



UNIVERSIDAD DE CUENCA

Facultad de Ingeniería

Carrera de Ingeniería de Sistemas

Implementación y evaluación de un dashboard para el análisis del
comportamiento de los estudiantes y predicción en Moodle

Trabajo de titulación previo a la obtención del título
de Ingeniero de Sistemas

Autores:

Sigua Loja Edison Fernando

CI: 0105874911

Correo electrónico: edisson.sigua1407@ucuenca.edu.ec

Aguilar Yaguana Bryan Alexander

CI: 1718570250

Correo electrónico: bryan.aguilar6174@gmail.com

Director:

Ing. Jorge Javier Maldonado Mahauad, PhD

CI: 1102959051

Cuenca, Ecuador

16-abril-2021



Resumen

El desarrollo tecnológico ha hecho que en las últimas dos décadas las Tecnologías de Información y Comunicación se involucren cada vez más en el proceso de enseñanza y traten de cambiar los modelos tradicionales de aprendizaje. Con el soporte de la tecnología moderna, se han desarrollado y perfeccionado plataformas que fomentan la adopción de un nuevo paradigma de aprendizaje virtual. Estas plataformas almacenan las interacciones de los estudiantes y los docentes con los recursos del curso en motores bases de datos, información que puede llegar a ser muy relevante, pero en muchos casos no ha sido procesada de forma que resulte útil para el uso de los docentes y estudiantes. Por ello, este estudio pretende implementar y evaluar un dashboard de análisis de comportamiento de los estudiantes y predicción de abandono en Moodle. La herramienta ayudará a los docentes a conocer lo que hacen los estudiantes antes, durante y después de una clase mediada por plataformas virtuales. Además, también ayudará a los estudiantes a gestionar su proceso de aprendizaje y monitorear de manera sencilla y eficaz su avance en el curso. Dada la naturaleza analítica de la investigación, se ha hecho uso de análisis exploratorios del modelo de datos de Moodle, evaluación de visualizaciones existentes y diseño de la herramienta basado en la arquitectura de Moodle para el desarrollo de un dashboard de visualizaciones y predicción de abandono.

Como resultado se implementó FlipMyLearning, un plugin para la plataforma Moodle que permite al docente monitorear el proceso de aprendizaje de los estudiantes para la toma de decisiones informadas. El plugin desarrollado contiene visualizaciones tanto para el docente como para el estudiante, divididas en diferentes secciones, cada una orientada a monitorear diferentes aspectos del curso.

La investigación realizada demuestra que las visualizaciones generadas resultaron útiles para los docentes y los estudiantes que participaron en el proceso de evaluación. Además, variables como el tiempo invertido, la cantidad de sesiones e indicadores relacionados a la profundidad cognitiva y la amplitud social son variables útiles para identificar grupos de estudiantes en riesgo de deserción.

Palabras claves: Analíticas del Aprendizaje. Dashboard. Moodle. Abandono. Predicción.



Abstract

The development of technology has meant that in the last two decades Information and Communication Technologies have become more and more involved in the teaching process and have tried to change traditional learning models. With the support of modern technology, platforms have been developed and perfected that encourage the adoption of a new virtual learning paradigm. These platforms store student and teacher interactions with course resources in database engines, information that can be very relevant, but in many cases has not been processed in a way that is useful for use by teachers and students. Therefore, this study aims to implement and evaluate a dashboard for student behavior analysis and dropout prediction in Moodle. The tool will help teachers to know what students do before, during and after a class mediated by virtual platforms. In addition, it will also help students manage their learning process and easily and effectively monitor their progress in the course. Given the analytical nature of the research, exploratory analysis of the Moodle data model, evaluation of existing visualizations and design of the tool based on Moodle architecture were used to develop a dashboard of visualizations and dropout prediction.

As a result, FlipMyLearning was implemented, a plugin for the Moodle platform that allows the teacher to monitor the learning process of students for informed decision making. The developed plugin contains visualizations for both the teacher and the student, divided into different sections, each oriented to monitor different aspects of the course.

The research conducted shows that the visualizations generated were useful for both teachers and students who participated in the evaluation process. In addition, variables such as time spent, number of sessions and indicators related to cognitive depth and social breadth are useful variables to identify groups of students at risk of dropping out.

Keywords: Learning Analytics. Dashboard. Moodle. Dropout. Prediction.



Índice

Capítulo 1: Introducción	23
1.1. Panorama general y justificación.....	23
1.2. Estado del Arte.....	25
1.2.1. Moodle y las Analíticas del Aprendizaje	25
1.2.2. Análisis del Comportamiento de los Estudiantes.....	27
1.3. Preguntas de investigación	29
1.4. Objetivos generales y específicos	29
1.5. Metodología.....	30
1.6. Estructura de la Tesis	33
1.7. Contribuciones de la Tesis.....	33
Capítulo 2: Tecnologías de la Información y Comunicación en la Educación.....	35
2.1. Aprendizaje Presencial	35
2.2. Aprendizaje Híbrido.....	35
2.3. Aprendizaje en Línea.....	36
2.4. Sistemas de Gestión de Aprendizaje	36
Capítulo 3: Sistema de Gestión de Aprendizaje Moodle.....	38
3.1. Moodle como EVEA.....	38
3.2. Características.....	39
3.2.1. Actividades	41
3.2.2. Recursos	42
3.2.3. Bloques.....	43
3.3. Arquitectura.....	44
3.4. Estructura de Directorios y Sistema	47
3.4.1. Código de Moodle.....	47
3.4.2. Datos de Moodle.....	49
3.4.3. Base de Datos de Moodle.....	50
Capítulo 4: Analíticas de Aprendizaje sobre Moodle	53
4.1. ¿Qué son las Analíticas de Aprendizaje?	53
4.2. Plugins de Analítica sobre Moodle	53
4.3. Patrones de Aprendizaje en Plataformas Virtuales.....	66
4.4. Predicción en Base a Interacciones de Aprendizaje	67
4.5. Analíticas de Aprendizaje y Dashboards.....	67
Capítulo 5: Conceptos Relacionado a Datos.....	69
5.1. Datos, Información y Conocimiento	69
5.1.1. Datos.....	69



5.1.2. Información.....	69
5.1.3. Conocimiento	70
5.2. Inteligencia de Negocios (<i>BI – Business Intelligence en inglés</i>).....	70
5.3. Data Warehouse	70
5.4. Data Warehousing	71
5.5. ETL (Extract, Transform and Load).....	71
5.6. Arquitectura del Data Warehousing.....	72
5.7. Modelos y Bases de Datos Multidimensionales	73
5.8. Metodología Hefesto.....	73
5.8.1. Análisis de requerimientos.....	74
5.8.2. Análisis de los OLTP	75
5.8.3. Modelo lógico del Datawarehouse.....	76
5.8.4. Integración de datos.....	79
5.9. Minería de Datos y Obtención del Conocimiento	80
5.9.1. Machine Learning.....	82
Capítulo 6: Desarrollo de un Dashboard de Visualización	87
6.1. Análisis Exploratorio de los Datos.....	87
6.1.1. Análisis de requerimientos.....	87
6.1.2. Análisis de los OLTP	95
6.2. Identificación de Patrones de Comportamiento.....	110
6.3. Detección de Variables de Predicción de Abandono	123
6.3.1. Profundidad Cognitiva	123
6.3.2. Amplitud social.....	125
6.3.3. Niveles de indicadores para Módulos de Actividad.....	126
6.3.4. Variables de Predicción del Modelo	127
6.3.5. Clustering.....	129
6.3.6. Métodos de división del tiempo	137
6.4. Planificación y Diseño del Dashboard.....	138
6.5. Implementación del Dashboard.....	146
6.5.1. Descripción general	146
6.5.2. Arquitectura.....	147
6.5.3. Tecnologías involucradas.....	148
6.5.4. Proceso de instalación.....	150
6.5.5. Funcionalidades	155
6.5.6. Funciones adicionales	196
6.5.7. Estructura de directorios.....	198



6.5.8. Proceso de desinstalación.....	199
6.6. Evaluación del Dashboard.....	201
6.6.1. Evaluación Vista Docente.....	202
6.6.2. Evaluación Vista Estudiante.....	207
Capítulo 7: Conclusiones.....	211
7.1. Conclusiones.....	211
7.2. Recomendaciones.....	215
7.3. Trabajos Futuros.....	216
Bibliografía.....	218
Anexos.....	231



Índice de Tablas

Tabla 1 <i>Descripción de las Actividades Disponibles en Moodle</i>	41
Tabla 2 <i>Descripción de los Recursos Disponibles en Moodle</i>	42
Tabla 3 <i>Descripción de los Distintos Tipos de Bloque en Moodle</i>	43
Tabla 4 <i>Descripción de los Módulos Principales en Moodle</i>	48
Tabla 5 <i>Descripción de las Tablas de Datos de Moodle Utilizadas para la Generación del Dashboard de Visualización</i>	51
Tabla 6 <i>Descripción de los Campos más Relevantes de la tabla mdl_logstore_standard_log</i>	96
Tabla 7 <i>Atributos analizados de la Tabla de Registro de Eventos de Moodle</i>	111
Tabla 8 <i>Log de Eventos con información detallada de la interacción</i>	111
Tabla 9 <i>Acciones Consideradas Relevantes</i>	112
Tabla 10 <i>Planificación en Semanas del Curso Planificación y Evaluación de Clases Virtuales</i> ...	113
Tabla 11 <i>Asignación de Niveles de Profundidad Cognitiva para las diferentes Actividades de Moodle</i>	124
Tabla 12 <i>Asignación de Niveles de Amplitud Social para las diferentes Actividades de Moodle</i> ...	126
Tabla 13 <i>Indicadores de Tiempo Invertido, Profundidad Cognitiva (CD) y Amplitud Social (SB)</i> 131	
Tabla 14 <i>Indicadores Normalizados de Tiempo Invertido, Profundidad Cognitiva (CD) y Amplitud Social (SB)</i>	131
Tabla 15 <i>indicadores más Relevantes que Caracterizan los Clústeres</i>	137
Tabla 16 <i>Rangos de Calificaciones Calculados para la Actividad denominada Proyecto 1</i>	176
Tabla 17 <i>Rangos de Calificaciones Calculados para la Actividad denominada Actividad 1.2</i>	176



Índice de Figuras

Figura 1 <i>Etapas de la Metodología empleada</i>	32
Figura 2 <i>Descripción General de la Arquitectura de Moodle</i>	45
Figura 3 <i>Interacción de los Elementos de la Arquitectura de Moodle</i>	45
Figura 4 <i>Capa de Moodle de la Arquitectura General</i>	46
Figura 5 <i>Áreas de Moodle</i>	47
Figura 6 <i>Lista de Directorios de Moodle 3.9</i>	48
Figura 7 <i>Categorías que Guardan Relación en la Base de Datos de Moodle 3.9</i>	51
Figura 8 <i>Descripción del Último Intento en una Evaluación/Cuestionario</i>	57
Figura 9 <i>Curva de Mejora en la Pestaña de Mi Progreso y Predicciones</i>	58
Figura 10 <i>Desempeño de los Compañeros en la Pestaña de Mi Progreso y Predicciones</i>	58
Figura 11 <i>Pregunta más Difícil en la Pestaña de Mi Progreso y Predicciones</i>	59
Figura 12 <i>Resumen de Intentos en la Pestaña de Mi Progreso y Predicciones</i>	59
Figura 13 <i>Categorías Desafiantes para Mí en la Pestaña de Análisis de Categoría de Preguntas</i> ..	60
Figura 14 <i>Calificaciones por Porcentajes en la Pestaña de Estado de Preguntas y Calificaciones</i> ..	61
Figura 15 <i>Análisis de Preguntas en la Pestaña de Estado de Preguntas y Calificaciones</i>	61
Figura 16 <i>Tabla de Calificaciones, Representación con Diagrama de Cajas y Bigotes</i>	63
Figura 17 <i>Representación Gráfica del Acceso al Contenido del Curso</i>	63
Figura 18 <i>Gráfico de Usuarios Activos</i>	64
Figura 19 <i>Cuadro de Envío de Tareas o Cuestionarios</i>	64
Figura 20 <i>Distribución de Visitas al Curso</i>	65
Figura 21 <i>Dashboard de IntelliBoard</i>	65
Figura 22 <i>Proceso de Extracción, Transformación y Carga de datos</i>	72
Figura 23 <i>Arquitectura del Data Warehousing</i>	72
Figura 24 <i>Estructura General de un Modelo Conceptual</i>	75
Figura 25 <i>Estructura General de un Modelo Conceptual Ampliado</i>	76
Figura 26 <i>Modelo de Tabla de Dimensión</i>	78
Figura 27 <i>Modelo de Tabla de Hechos</i>	79
Figura 28 <i>Representación de Uniones para Tablas de Hechos y Dimensiones</i>	79
Figura 29 <i>Elbow curve, para el cual en $k = 3$ será el número de clusters óptimo</i>	85
Figura 30 <i>Silhouette Curve, para el cual en $k = 4$ será el número de clusters óptimo</i>	86
Figura 31 <i>Mapas de Procesos de la Semana 0</i>	115
Figura 32 <i>Mapas de Procesos de la Semana 1</i>	116



Figura 33 Mapas de Procesos de la Semana 2	118
Figura 34 Mapas de Procesos de la Semana 3	119
Figura 35 Mapas de Procesos de la Semana 4	121
Figura 36 Mapas de Procesos de la Semana 5	122
Figura 37 Matriz de Profundidad Cognitiva y Amplitud Social de Actividades y Recursos de Moodle	126
Figura 38 Curva del Codo	132
Figura 39 Gráficas del Método de la Silueta para $k = 3$	133
Figura 40 Gráfico de Distribución Datos con Reducción de Componentes	135
Figura 41 Gráfico de Distribución Datos Agrupados en tres Clústeres	136
Figura 42 Prototipo para conocer la cantidad de visitas realizadas al curso por (día de la semana) / (hora del día) a modo de mapa de calor	140
Figura 43 Prototipo para visualizar un reporte de visitas, tiempo y progreso de los estudiantes en el curso a modo de tabla interactiva	141
Figura 44 Prototipo para visualizar la distribución de sesiones por diferentes rangos de tiempo mediante un gráfico de líneas	142
Figura 45 Prototipo para conocer el tiempo promedio invertido a comparación del tiempo promedio planificado por el docente	143
Figura 46 Prototipo para conocer los Diferentes Estados de envío de Actividades para su Calificación	144
Figura 47 Prototipo para conocer la Distribución de Acceso a los Recursos por parte de los Estudiantes	144
Figura 48 Prototipo para conocer el Promedio de Calificaciones de las Diferentes Actividades Calificables	145
Figura 49 Prototipo para conocer la Distribución de Intentos para cada una de las Evaluaciones	146
Figura 50 Arquitectura General	148
Figura 51 Extracción del Directorio fliplearning	151
Figura 52 Directorio local en el Directorio de Instalación de Moodle	151
Figura 53 fliplearning en el directorio local de Moodle	152
Figura 54 Sesión de administrador iniciada en la sección de Área Personal del sitio	152
Figura 55 Comprobación de requisitos y confirmación de instalación para FlipMyLearning	153
Figura 56 Confirmación de Actualización de la Base de Datos de Moodle	154



Figura 57 <i>Confirmación de Instalación Exitosa</i>	154
Figura 58 <i>Ubicación de FlipMyLearning en Moodle</i>	155
Figura 59 <i>Ícono de la Herramienta FlipMyLearning</i>	156
Figura 60 <i>Opción Inicial de Configuración de Semanas de Estudio</i>	156
Figura 61 <i>Configuración de Semanas de Estudio</i>	157
Figura 62 <i>Advertencias de FlipMyLearning al no Cumplir una de las Condiciones de Configuración establecidas</i>	158
Figura 63 <i>Secciones Disponibles de FlipMyLearning después de la Configuración de Semanas</i> ..	158
Figura 64 <i>Sección de Ayuda para Secciones de FlipMyLearning</i>	159
Figura 65 <i>Modal de Ayuda para las Diferentes Secciones de FlipMyLearning</i>	159
Figura 66 <i>Función Acerca de este gráfico para las Diferentes Visualizaciones de FlipMyLearning</i>	160
Figura 67 <i>Indicadores Generales del Curso</i>	161
Figura 68 <i>Recursos por Semanas de Estudio</i>	162
Figura 69 <i>Mapa de calor de Sesiones por Semana</i>	163
Figura 70 <i>Tabla Interactiva para Evidenciar un Reporte de Participación de los Estudiantes</i>	164
Figura 71 <i>Gráfico de barras para evidenciar el Tiempo promedio invertido por los Estudiantes a Comparación del Tiempo promedio planificado por el Docente a lo largo de una semana de estudio</i>	166
Figura 72 <i>Mapa de calor de Sesiones por Día y Hora</i>	167
Figura 73 <i>Gráfico de Líneas para evidenciar la Cantidad de Sesiones en Diferentes intervalos de tiempo a lo largo de una semana de estudio</i>	168
Figura 74 <i>Histograma para Representar la Cantidad de Estudiantes que han enviado, que no han enviado o han enviado tarde una Tarea</i>	169
Figura 75 <i>Estructura del Modal de Envío de Correo Electrónico a Estudiantes</i>	170
Figura 76 <i>Envío de Correo Electrónico</i>	170
Figura 77 <i>Correo Electrónico Generado por FlipMyLearning</i>	171
Figura 78 <i>Gráfico de barras para Evidenciar el Acceso a Recursos y Actividades de los estudiantes</i>	172
Figura 79 <i>Selector de Unidades Definidas para Actividades Evaluables</i>	173
Figura 80 <i>Promedio de Actividades Evaluables</i>	173
Figura 81 <i>Ejemplo de actividad con ponderación máxima de 35 puntos</i>	175
Figura 82 <i>Ejemplo de Actividad con Ponderación máxima de 100 puntos</i>	175



Figura 83 Rangos de Calificación de una Actividad con Ponderación Máxima de 35 puntos.....	177
Figura 84 Rangos de Calificación de una Actividad con Ponderación máxima de 100 puntos	177
Figura 85 Selector de Evaluaciones en una Semana de Estudio.....	178
Figura 86 Histograma de Intentos de Preguntas para una evaluación	179
Figura 87 Detalle de la Pregunta en la Evaluación	180
Figura 88 Nivel de Dificultad de cada una de las Preguntas de una Evaluación.....	181
Figura 89 Selector de Grupos de Estudiantes Identificados.....	181
Figura 90 Tabla Interactiva de Estudiantes del Grupo	182
Figura 91 Información del Perfil del Estudiante.....	183
Figura 92 Recursos divididos por etiquetas de Accedidos, Completos y Total.	184
Figura 93 Recursos del Curso.	185
Figura 94 Recursos de las diferentes Semanas de Estudio	186
Figura 95 Evolución de las Sesiones y Tiempo Invertido de un Estudiante	187
Figura 96 Acercamiento de Evolución de las Sesiones y Tiempo Invertido de un Estudiante en un rango de Tiempo.....	188
Figura 97 Comparación de Calificaciones en las diferentes Actividades Evaluables.....	189
Figura 98 Información del Perfil del Estudiante.....	189
Figura 99 Recursos divididos por etiquetas de Accedidos, Completos y Total	190
Figura 100 Mapa de calor de Sesiones por Semana.....	191
Figura 101 Evolución de las Sesiones y Tiempo Invertido de un Estudiante.....	192
Figura 102 Comparación de Calificaciones en las diferentes Actividades Evaluables	193
Figura 103 Gráfico de barras para evidenciar el Tiempo promedio invertido del Estudiante a Comparación del Tiempo promedio planificado por el Docente a lo largo de una semana de estudio	194
Figura 104 Mapa de calor de Sesiones por Día y Hora	195
Figura 105 Histograma Apilado de Interacción por Tipos de Recursos.....	196
Figura 106 Usar Filtro de Grupos del Curso	197
Figura 107 Menú de Opciones Adicionales para los Gráficos de FlipMyLearning.....	198
Figura 108 Estructura de directorios de FlipMyLearning.....	198
Figura 109 Extensiones locales de Moodle.....	199
Figura 110 Lista de extensiones locales Disponibles en Moodle	200
Figura 111 Proceso de Desinstalación de FlipMyLearning.....	200



Figura 112 <i>Criterios de Usabilidad para la Evaluación de las Visualizaciones de la Sección Indicadores Generales (Vista Docente)</i>	202
Figura 113 <i>Criterios de Usabilidad para la Evaluación de las Visualizaciones de la Sección Sesiones de Estudio (Vista Docente)</i>	203
Figura 114 <i>Criterios de Usabilidad para la Evaluación de las Visualizaciones de la Sección Seguimiento de Calificaciones (Vista Docente)</i>	203
Figura 115 <i>Criterios de Usabilidad para la Evaluación de las Visualizaciones de la Sección Seguimiento de Tareas (Vista Docente)</i>	204
Figura 116 <i>Criterios de Usabilidad para la Evaluación de las Visualizaciones de la Sección Seguimiento de Evaluaciones (Vista Docente)</i>	205
Figura 117 <i>Criterios de Usabilidad para la Evaluación de las Visualizaciones de la Sección Deserción (Vista Docente)</i>	205
Figura 118 <i>Criterios de Usabilidad para la Evaluación de la Herramienta en General (Vista Docente)</i>	206
Figura 119 <i>Criterios de Usabilidad para la Evaluación de las Visualizaciones de la Sección Indicadores Generales (Vista Estudiante)</i>	208
Figura 120 <i>Criterios de Usabilidad para la Evaluación de las Visualizaciones de la Sección Sesiones de Estudio (Vista Estudiante)</i>	208
Figura 121 <i>Criterios de Usabilidad para la Evaluación de la Herramienta en General (Vista Estudiante)</i>	209

Índice de Anexos

Anexo 1. Certificado de presentación del artículo de posición.....	231
Anexo 2. Relaciones de tablas de la base de datos de Moodle, identificando tablas relevantes y atributos de interés para el estudio.....	233



Listado de Abreviaturas

- **AI** – *Artificial Intelligence*: Inteligencia artificial.
- **ICT** – *Information and Communications Technology*: Tecnologías de información y comunicación.
- **LA** – *Learning Analytics*: Analíticas del Aprendizaje.
- **EDM** – *Educational Data Mining*: Minería de Datos Educativos.
- **B-learning** – *Blended Learning*: Aprendizaje Combinado.
- **WBT** – *Web Based Training*: Entrenamiento Basado en la Web.
- **LMS** – *Learning Management System*: Sistemas de Gestión de Aprendizaje.
- **LCMS** – *Learning Content Management System*: Sistemas de Gestión de Contenidos de Aprendizaje.
- **CMS** – *Course Management System*: Sistemas Manejadores de Cursos.
- **Moodle** – *Modular Object-Oriented Dynamic Learning Environment*: Entorno de Aprendizaje Dinámico Orientado a Objetos Modulares.
- **MOOC** – *Massive Open Online Course en inglés*: Cursos en línea masivos y abiertos.
- **SRL** – *Self-Regulated Learning*: Aprendizaje Autorregulado.
- **ML** – *Machine Learning*: Aprendizaje automatizado.
- **ETL** – *Extract, Transform and Load*: Extracción, Transformación y Carga.
- **DBR** – *Design Based Research*: Metodología de Investigación basada en el diseño.
- **E-learning** – *Electronic Learning*: Aprendizaje Electrónico.
- **MLE** – *Managed Learning Environment*: Entorno de Aprendizaje Gestionado.
- **LSS** – *Learning Support System*: Sistema de Apoyo al Aprendizaje.
- **LP** – *Learning Platform*: Plataforma de Aprendizaje.
- **VLE** – *Virtual Learning Environment*: Entornos de Aprendizaje Virtual.
- **EVEA**: Entorno Virtual de Enseñanza y Aprendizaje.
- **FC** – *Flipped Classroom*: Aula invertida.
- **LTI** – *Learning Tools Interoperability*: Herramientas de Aprendizaje Interoperables.
- **SCORM** – *Sharable Content Object Reference Model*: Modelo Referenciado de Objetos de Contenido Compartible.
- **COLLES** – *Constructivist On-Line Learning Environment Survey*: Encuesta Constructivista sobre el Entorno de Aprendizaje en Línea.



- **ATTLS** – *Attitudes to Thinking and Learning Survey*: Encuesta de Actitudes hacia el Pensamiento y el Aprendizaje.
- **LAMP**: Infraestructura de internet basada en Linux, Apache, MySQL y PHP.
- **KPI** – *Key Performance Indicator*: Indicador Clave de Rendimiento.
- **OLTP** – *OnLine Transaction Processing*: Sistema de Procesamiento de Transacciones en Línea.
- **SGBD** – *Database Management System*: Sistema Gestor de Base de Datos.
- **ITS** – *Intelligent Tutoring System*: Sistemas de Tutoría Inteligente.
- **DM** – *Data Mining*: Minería de Datos.
- **MVC** – *Model View and Controller*: Modelo Vista Controlador.
- **AJAX** – *Asynchronous JavaScript and XML*: JavaScript Asíncrono y XML.
- **API** – *Application Programming Interface*: Interfaz de Programación de Aplicaciones.
- **JSON** – *JavaScript Object Notation*: Notación de Objetos de JavaScript.
- **BI** – *Business Intelligence*: Inteligencia de Negocios.
- **RDBMS** – *Relational Database Management System*: Sistema de Gestión de Bases de Datos Relacionales.
- **UTC** – *Coordinated Universal Time*: Tiempo Universal Coordinado
- **CD** – *Cognitive Depth*: Profundidad Cognitiva
- **SB** – *Social Breadth*: Amplitud Social



Cláusula de Propiedad Intelectual

Bryan Alexander Aguilar Yaguana, autor del trabajo de titulación "Implementación y evaluación de un dashboard para el análisis del comportamiento de los estudiantes y predicción en Moodle", certifico que todas las ideas, opiniones y contenidos expuestos en la presente investigación son de exclusiva responsabilidad de su autor.

Cuenca, 24 de marzo de 2021

Bryan Alexander Aguilar Yaguana

C.I: 1718570250



Cláusula de Propiedad Intelectual

Edisson Fernando Sigua Loja, autor del trabajo de titulación "Implementación y evaluación de un dashboard para el análisis del comportamiento de los estudiantes y predicción en Moodle", certifico que todas las ideas, opiniones y contenidos expuestos en la presente investigación son de exclusiva responsabilidad de su autor.

Cuenca, 24 de marzo de 2021

Edisson Fernando Sigua Loja

C.I: 0105874911



Cláusula de licencia y autorización para publicación en el Repositorio Institucional

Bryan Alexander Aguilar Yaguana en calidad de autor y titular de los derechos morales y patrimoniales del trabajo de titulación "Implementación y evaluación de un dashboard para el análisis del comportamiento de los estudiantes y predicción en Moodle", de conformidad con el Art. 114 del CÓDIGO ORGÁNICO DE LA ECONOMÍA SOCIAL DE LOS CONOCIMIENTOS, CREATIVIDAD E INNOVACIÓN reconozco a favor de la Universidad de Cuenca una licencia gratuita, intransferible y no exclusiva para el uso no comercial de la obra, con fines estrictamente académicos.

Asimismo, autorizo a la Universidad de Cuenca para que realice la publicación de este trabajo de titulación en el repositorio institucional, de conformidad a lo dispuesto en el Art. 144 de la Ley Orgánica de Educación Superior.

Cuenca, 24 de marzo de 2021

Bryan Alexander Aguilar Yaguana

C.I: 1718570250



Cláusula de licencia y autorización para publicación en el Repositorio Institucional

Edisson Fernando Sigua Loja en calidad de autor y titular de los derechos morales y patrimoniales del trabajo de titulación "Implementación y evaluación de un dashboard para el análisis del comportamiento de los estudiantes y predicción en Moodle", de conformidad con el Art. 114 del CÓDIGO ORGÁNICO DE LA ECONOMÍA SOCIAL DE LOS CONOCIMIENTOS, CREATIVIDAD E INNOVACIÓN reconozco a favor de la Universidad de Cuenca una licencia gratuita, intransferible y no exclusiva para el uso no comercial de la obra, con fines estrictamente académicos.

Asimismo, autorizo a la Universidad de Cuenca para que realice la publicación de este trabajo de titulación en el repositorio institucional, de conformidad a lo dispuesto en el Art. 144 de la Ley Orgánica de Educación Superior.

Cuenca, 24 de marzo de 2021

Edisson Fernando Sigua Loja

C.I: 0105874911



Agradecimiento

Agradezco a mi familia que durante todo el trayecto de la universidad fueron mi principal apoyo. En especial a mis padres que me dieron la fortaleza para ser mejor cada día y fueron mi principal motivación.

A todos los docentes a los que pude conocer en la universidad y contribuyeron con sus conocimientos y enseñanzas para que consiguiera mi título profesional. De manera especial agradezco al Ingeniero Jorge Maldonado por el tiempo dedicado y sus guías para la realización de nuestro trabajo de titulación.

A todos los amigos con los que pude compartir en la universidad y que sumaron en mi vida. En especial a Bryan, Esteban, Paola, Christian, Freddy, Paúl, Ariel con quienes tengo gratos recuerdos de los últimos años de la universidad.

Edisson Sigua L.



Agradecimiento

Muy orgulloso de poder hacer mención de todos a quienes agradezco el poder llegar al cumplimiento de esta meta tan anhelada que es mi título profesional a través de este trabajo de tesis. Deseo empezar mi agradecimiento con lo siguiente:

A Dios por darme la voluntad y la fuerza para seguir adelante en todo este proceso de realización personal.

A mi familia por ser un apoyo constante en todo este proceso. Especialmente a mis padres y hermanos quienes me han apoyado cada día a pesar de la distancia y cualquier circunstancia.

A mi director de tesis, Ingeniero Jorge Maldonado, por su valiosa dirección y apoyo, por su paciencia, profesionalismo, críticas constructivas, experiencia y educación. Factores que han sido mi fuente de motivación durante este tiempo.

A cada uno de los docentes, quienes supieron transmitir sus conocimientos y experiencias a través de las aulas universitarias, que de seguro serán herramientas útiles en el ámbito profesional.

A mis amigos más cercanos, Fercho, Christian, Freddy, Esteban, Ariel, Pao, Gaby y William por brindarme su apoyo, compartir grandes momentos, experiencias inolvidables y hacer que la vida universitaria sea de las mejores etapas de mi vida.

Finalmente, agradezco a toda la comunidad universitaria: grupos estudiantiles, asos escuelas, grupos académicos, etc. Me siento plenamente orgulloso en ser parte de esta gran institución.

Bryan Aguilar Y.



Dedicatoria

Este trabajo es la conclusión de años de constante esfuerzo y dedicación. Por lo que dedico el mismo:

A mi madre y mi padre, que estuvieron conmigo en todos los momentos positivos y negativos de mi vida. Sin su apoyo, esto no hubiese sido posible.

A mis hermanos, con los cuales aprendí mucho de lo que me ha llevado la persona que soy, y me siento inmensamente orgulloso de poder ser un ejemplo para ellos.

A mis demás familiares, que supieron apoyarme en los momentos en los que más necesité.

A todos mis compañeros de carrera y a mis amigos de la aso sistemas por brindarme su valiosa amistad. Espero de corazón que cumplan todas sus metas.

Edisson Sigua L.



Dedicatoria

Dedico este trabajo y esfuerzo en primer lugar a Dios y a la Virgen, quienes me dan la fe y voluntad que necesito cada día para ser mejor y no rendirme en este camino lleno de sueños y desafíos llamado vida.

A toda mi familia, quienes han sido un pilar fundamental durante toda mi vida y más aún en mi desarrollo personal y profesional. Por brindarme consejos, compartir experiencias, y sobre todo gratos y buenos momentos. Especialmente a mis padres quienes con cariño y apoyo incondicional supieron guiarme y enseñarme grandes valores; muchos de mis logros se los debo a ustedes. Espero puedan sentirse orgullosos de mí como yo lo estoy de ser su hijo.

A mi hermano y hermana para que vean en mí un ejemplo a seguir y sepan que, con trabajo duro, dedicación y esfuerzo, ninguna meta es imposible de cumplir. ¡Se lo dedico a ustedes hermanos!

Bryan Aguilar Y.



Capítulo 1: Introducción

1.1. Panorama general y justificación

El desarrollo tecnológico ha sido uno de los pilares fundamentales en la evolución humana hacia lo que somos hoy en día. La razón principal de esta relación de desarrollo mutuo se basa en el hecho de que todo proceso evolutivo es una combinación de componentes y/o procesos para cubrir algún propósito en un área humana específica (Coccia & Watts, 2019). Algunas de las áreas que más se han beneficiado del desarrollo tecnológico son la computación, medicina, robótica, electricidad y telecomunicaciones (Temple, 2020). Actualmente, es posible establecer una conversación en tiempo real con otra persona en cualquier parte del mundo y por medio de las redes sociales se puede acceder a información de cualquier acontecimiento en instantes. Con los avances logrados en Inteligencia Artificial (*AI – Artificial Intelligence en inglés*) se han empezado a construir vehículos con sistemas de manejo autónomos capaces de llegar a un destino sin necesidad de interacción humana (Teoh & Kidd, 2017). También gracias a la tecnología, la medicina ha tenido grandes avances como el mapeo del Genoma Humano, el mismo que se pudo estudiar sistemáticamente con la ayuda de programas informáticos para la obtención de patrones y asociaciones (Zwart, 2015). Además, a corto plazo, el número de dispositivos conectados a Internet aumentará de manera exponencial gracias a la velocidad de transferencia rápida, alta confiabilidad, seguridad robusta, bajo consumo de energía y gran cantidad de conexiones de las redes 5G (Hui, y otros, 2020). De esta forma es posible ver cómo la tecnología tiene influencia en todas las áreas humanas.

En el caso específico del área de la educación, los avances en las Tecnologías de Información y Comunicación (*ICT – Information and Communications Technology en inglés*) han tratado de contribuir al enfoque tradicional de aprendizaje presencial, convirtiéndose en parte integral de los procesos de enseñanza y aprendizaje (Juhaňák, Zounek, & Rohlíková, 2019) y generando áreas de interés investigativo tales como las Analíticas del Aprendizaje (*LA – Learning Analytics en inglés*) y Minería de Datos Educativos (*EDM – Educational Data Mining en inglés*) (Den Beemt, Buijs, & Der Aalst, 2018). Gracias al uso de las tecnologías, por medio de la web y otros canales, el sector educativo ha podido superar las limitaciones geográficas/temporales que los esquemas tradicionales de enseñanza-aprendizaje conllevan, dando lugar así al nacimiento de nuevos modelos de enseñanza y aprendizaje como e-formación, e-learning o b-learning (García Alba, 2010). De los tres modelos antes mencionados, el que incorpora significativamente las ICT a su proceso de enseñanza es el Aprendizaje Combinado (*b-learning – Blended Learning en inglés*), el cual se define como un enfoque que combina los beneficios ofrecidos por los componentes de aprendizaje presencial y en



línea (Abubakar, Kamsin, & Abdullah, 2020). Como resultado, en los últimos años ha aumentado el interés en el diseño de entornos de aprendizaje híbridos, ya que combinar las actividades de enseñanza presencial y en línea ofrece nuevas oportunidades para potenciar y optimizar el aprendizaje (Spanjers, y otros, 2015). De hecho, el aprendizaje híbrido, respaldado por plataformas virtuales, se está convirtiendo en una de las perspectivas de aprendizaje y enseñanza más beneficiosas para la educación superior (Gutiérrez Braojos, Montejo Gamez, Marin Jimenez, & Campaña, 2018).

B-learning es un modelo de aprendizaje híbrido apoyado por plataformas virtuales, las cuales tienen sus orígenes en el Entrenamiento Basado en la Web (*WBT – Web Based Training en inglés*). La evolución de WBT ha propiciado el surgimiento de Sistemas de Gestión de Aprendizaje (*LMS – Learning Management System en inglés*) y sus variantes; los Sistemas de Gestión de Contenidos de Aprendizaje (*LCMS – Learning Content Management System en inglés*) y los Sistemas Manejadores de Cursos (*CMS – Course Management System en inglés*) (García Alba, 2010). Estos sistemas se usan para mediar el proceso de aprendizaje de los estudiantes en ambientes híbridos de (Codish, Rabin, & Ravid, 2019).

La información en las plataformas LMS se produce de las interacciones de los estudiantes con la plataforma y se almacena en forma de registros generados automáticamente (Juhaňák, Zounek, & Rohlíková, 2019). A pesar de las capacidades que estas plataformas aportan al proceso de aprendizaje, éstas se encuentran infravaloradas debido a que las interacciones de los estudiantes en estas se han limitado a la descarga de material de clase, la entrega de tareas y la lectura de anuncios (Costa, Alvelos, & Teixeira, 2012). La subutilización de las plataformas LMS se debe en gran medida al hecho de que por muchos años estos sistemas han representado una especie de caja negra que únicamente almacena registros de interacción. Sin embargo, con el surgimiento de áreas de investigación como LA y EDM se les ha dado más atención en busca de explorar el potencial de los datos que tienen almacenados estos sistemas (Juhaňák, Zounek, & Rohlíková, 2019).

Actualmente, la falta de herramientas en los LMS para apoyar a los profesores es una de las razones por las cuales éstos se han convertido en repositorios de material educativo más que en verdaderos sistemas que ayuden a gestionar el aprendizaje. En este contexto, resulta prioritario el desarrollar una herramienta que permita a los docentes monitorear y visualizar el comportamiento de los estudiantes en la plataforma (Juhaňák, Zounek, & Rohlíková, 2019) a fin de poder tener evidencia de su proceso de aprendizaje en diferentes momentos de la clase (esto es antes, durante, después). Por otro lado, esta información también podría contribuir proporcionando información al sobre la calificación final del estudiante en el curso de manera anticipada (ya que, (G. de Barba, y otros, 2020) sugiere que el rendimiento académico guarda relación con el número de sesiones en la plataforma



virtual) y realizar intervenciones oportunas que permitan mitigar el riesgo de abandono temprano. Además, conocer los itinerarios de aprendizaje que realizan los estudiantes, puede servir para realizar un análisis sobre el diseño instruccional del curso y se podrían utilizar los resultados para volver a diseñar cursos existentes (Den Beemt, Buijs, & Der Aalst, 2018), mejorar los procesos de aprendizaje y predecir la deserción estudiantil (Moreno Marcos, Muñoz Merino, Alario Hoyos, Estévez Ayres, & Delgado Kloos, 2018).

En este trabajo de tesis, se propone la implementación y evaluación de un dashboard de análisis de comportamiento y predicción de abandono de los estudiantes en la plataforma Moodle que permita tanto a docentes como a estudiantes monitorear el proceso de aprendizaje en un curso online. La herramienta llamada FlipMyLearning proveerá a los docentes la capacidad de saber qué realmente está sucediendo en el transcurso de un curso de manera que sea posible implementar cambios tempranos para alinear los comportamientos estudiantiles a los objetivos académicos. Asimismo, será posible conocer los diferentes tipos de comportamientos de aprendizaje existentes en un curso, de manera que el docente tenga la capacidad de distinguir a los estudiantes comprometidos con el aprendizaje y tomar correctivos con aquellos cuyo nivel de compromiso no es el adecuado. Además, se implementará un proceso de evaluación del dashboard para verificar su funcionamiento y utilidad en un contexto real, de manera que se genere una retroalimentación que ayude a identificar debilidades no contempladas y nuevas necesidades que podrían servir para estudios futuros.

1.2. Estado del Arte

1.2.1. Moodle y las Analíticas del Aprendizaje

La última década ha sido testigo de la rápida adopción y el uso de varios LMS en la educación superior (Mwalumbwe & Mtebe, 2017). Entre las más populares están Moodle, edX, Coursera, Claroline, Absorb LMS, SAP Litmos, Docebo, entre otros (Anthony, 2018). A pesar de la continua adopción y uso de estos LMS, hay poca evidencia que sugiera que en realidad esté mejorando el rendimiento de aprendizaje de los estudiantes (Einhardt, Aires Tavares, & Cechinel, 2016).

Moodle, Entorno de Aprendizaje Dinámico Orientado a Objetos Modulares (*Moodle – Modular Object-Oriented Dynamic Learning Environment en inglés*) es una de las plataformas más populares debido a su licencia de software libre, facilidad de personalización y arquitectura modular (García Alba, 2010). Moodle permite a los docentes crear cursos en línea con un enfoque en la interacción y la construcción colaborativa de contenido, y está en continua evolución (Luna, Castro, & Romero, 2017). Moodle recopila grandes cantidades de información sobre las interacciones de los



estudiantes. Esta plataforma proporciona a los docentes retroalimentación de los cursos que dirigen a manera de módulos que aportan funcionalidades visualización de bitácoras, estadísticas, monitoreo de eventos y diversos tipos de reportes, entre ellos de desglose de competencias, de actividad, de participación, entre otros (Yassine, Kadry, & Sicilia, 2016). Muchas de estas herramientas proporcionadas por Moodle ofrecen información y capacidades de filtrado, sin embargo, los datos que proporcionan no aportan información significativa sobre el curso (García Alba, 2010). Extraer este tipo de datos útiles y transformar dicha información en conocimiento procesable es una tarea difícil (Luna, Castro, & Romero, 2017).

EDM y LA ofrecen perspectivas, metodologías, técnicas y herramientas diferentes con el objetivo de facilitar el proceso de descubrimiento de conocimiento a partir de datos educativos, teniendo en cuenta procesos que reflejen el comportamiento de los estudiantes. El propósito de LA y EDM es comprender y optimizar el proceso de aprendizaje y los entornos en los que ocurre (Shaun Bake & Salvador Inventado, 2014). Romero, Ventura y García (2008) fueron de los primeros autores en realizar una investigación de cómo integrar y combinar técnicas de extracción y procesamiento de datos como estadísticas, visualización, clasificación, agrupación, reglas de asociación y detección de patrones sobre datos producidos por la plataforma Moodle.

Uno de los principales estudios que deja ver la utilidad de EDM y LA en el descubrimiento de información a partir de datos educativos en el LMS Moodle es (Juhaňák, Zounek, & Rohlíková, 2019), en el cual se exploraron patrones de interacción de estudiantes y se extrajeron cuatro tipos de comportamiento que adoptan los estudiantes al rendir pruebas/exámenes. Sin embargo, a pesar de que este estudio hace una contribución sustancial a la investigación existente de análisis del aprendizaje, sus resultados no han podido ser explotados por la falta de una herramienta que permita visualizar los patrones de comportamiento hallados. Además, el estudio se enfoca en la detección de patrones de comportamiento en las evaluaciones a los estudiantes, por lo que los tipos de comportamiento detectables están restringidos a este contexto, sin tomar en cuenta las interacciones con los demás materiales del curso.

Por otro lado, existe también una gran cantidad de plugins de terceros que pueden ser instalados para extender las funcionalidades nativas de Moodle relacionado a nuevos tipos de actividades, preguntas para exámenes, reportes, integración con otros sistemas, entre otras capacidades (Romero, Ventura, & Garcia, 2008). Además de ello, a partir del lanzamiento de la versión 3.4 de Moodle, se ha integrado un módulo de LA que permite realizar análisis descriptivo y predictivo del rendimiento de los estudiantes basado en los datos almacenados por la plataforma y con el objetivo de tomar acciones correctivas de manera temprana (Moodle, 2020).



A pesar de todas las funcionalidades que proporciona Moodle al monitoreo del proceso de aprendizaje de los estudiantes, es evidente que aún es necesario la implementación de analíticas en diferentes grados de granularidad que involucren la mayor cantidad posible de interacciones del alumno en el curso con el afán de realizar una mejor monitorización del proceso de aprendizaje (Zorrilla & García Saiz, 2013). También es visible que no se ha explotado todo el poder de los datos que almacena Moodle relacionados a la interacción de los estudiantes con cada tipo de recursos de un curso y a la cantidad de sesiones que realiza un estudiante (G. de Barba, y otros, 2020), datos que guardan mucha información referente al compromiso del estudiante con la aprobación de un curso (Juhaňák, Zounek, & Rohlíková, 2019). Además de ello, un inconveniente que presenta la implementación del análisis descriptivo y predictivo que ofrece Moodle es su enfoque únicamente en los datos del estudiante en un curso. En este sentido, las interacciones fuera del aula de clase más la información del estudiante podría ser útil para mejorar el análisis descriptivo y predictivo implementado por Moodle.

1.2.2. Análisis del Comportamiento de los Estudiantes

Si bien existen ciertos tipos de analíticas en Moodle como las mencionadas anteriormente, existe también la necesidad de poder extraer y analizar las interacciones que tienen los estudiantes y su comportamiento dentro de la plataforma, esto con el objetivo de evidenciar cuáles son sus rutas de aprendizaje, su compromiso y su rendimiento académico.

En este contexto, Van Den Beemt realizó una contribución a la investigación del análisis del aprendizaje en plataformas virtuales con su estudio (Den Beemt, Buijs, & Der Aalst, 2018). El propósito de su trabajo fue obtener información sobre cómo los estudiantes aprobados y reprobados distribuyen sus actividades de manera diferente a lo largo de las semanas de un curso. Para ello, los autores exploraron la relación entre el comportamiento de aprendizaje y el progreso del aprendizaje en un curso MOOC (*MOOC – Massive Open Online Course en inglés*) con un enfoque basado en minería de procesos y desde la perspectiva del constructivismo personal para analizar secuencias de observación de videos y rendición de pruebas de los estudiantes. Como resultado, los autores concluyen que la minería de procesos, combinada con estadísticas tradicionales y aplicada desde una perspectiva de constructivismo personal muestra ser un enfoque fructífero para investigar el comportamiento y el progreso del aprendizaje, por lo que puede usarse junto con la variable del tiempo para describir cómo se ordenan las actividades de los estudiantes en patrones y a qué resultados conducen estas secuencias.



Por otro lado, una de las investigaciones más actuales en el área del análisis del aprendizaje es (Valle Torre, Tan, & Hauff, 2020). Este estudio presenta una herramienta de análisis de registros en la plataforma MOOC edX que está basada en un navegador web. La herramienta permite la generación de unidades denominadas sesiones de aprendizaje, en las cuales se visualizan tres aspectos principales. Primero, rutas de aprendizaje: las rutas que los estudiantes toman a través de los componentes MOOCS. Segundo, participación en foros: cómo se desarrollan los números de los carteles en foros del curso. Tercero, secuencias de observación de videos: determinan en qué medida, los estudiantes que aprueban siguen rutas de aprendizaje observando videos. El objetivo del estudio es equipar a los investigadores con una herramienta de análisis de datos y capacitar a aquellos con menos experiencia en el área, con el afán de que el análisis de grandes conjuntos de datos MOOC se realice de manera efectiva y eficiente.

Otro estudio reciente es De Barba, y otros, 2020. En este se examinan los datos de comportamiento de 9,272 estudiantes en las sesiones de un MOOC y la relación con su compromiso, calificación y datos auto informados que miden aspectos del Aprendizaje Autorregulado (*SRL – Self-Regulated Learning en inglés*). La investigación se realiza mediante un enfoque exploratorio temporal en relación con aspectos macro (distribución de sesiones en relación con su duración y frecuencia a lo largo del curso) y micro (actividades dentro de sesiones). El principal resultado de este estudio fue determinar cómo la participación y el logro de un estudiante en un MOOC se relaciona con la distribución de las sesiones y actividades de dichas sesiones. Además, se determinó que los estudiantes exitosos tuvieron sesiones más frecuentes y largas en todo el curso. Este estudio contribuye al campo de la investigación de tecnología educativa con el uso de un enfoque temporal como una forma de proporcionar una interpretación significativa del comportamiento de los estudiantes.

El análisis de comportamiento de los estudiantes sobre MOOC también ha sido testigo de una proliferación de estudios empíricos sobre la tasa de deserción de los alumnos, los patrones de participación y la autorregulación (Valle Torre, Tan, & Hauff, 2020). Estas áreas de investigación a menudo exigen más que simples análisis estadísticos de encuestas o medidas de autoinforme para sacar conclusiones sobre la experiencia de aprendizaje en línea y el compromiso de aprendizaje de los estudiantes (Bogarín, Cerezo, & Romero, A survey on educational process mining, 2017). Dashboards de aprendizaje y modelos predictivos constituyen dos de las principales soluciones a este desafío (Abubakar, Kamsin, & Abdullah, 2020). Por una parte, los dashboard de aprendizaje pueden mostrar datos a través de diferentes visualizaciones, como gráficos, indicadores, y mapas (Schwendimann, Rodríguez Triana, Vozniuk, Prieto, & Shirvani Boroujeni, 2017). Por otro lado, los



modelos predictivos brindan información oportuna sobre los alumnos en riesgo de deserción con el objetivo de generar intervenciones oportunas por parte del docente (Moreno Marcos, y otros, 2020). Si bien estas soluciones están creciendo en popularidad en los últimos años, la incógnita de cómo proveer, analizar y presentar esta información de manera significativa a los diferentes interesados sigue sin resolverse (Abubakar, Kamsin, & Abdullah, 2020).

1.3. Preguntas de investigación

A pesar de todos los esfuerzos realizados para monitorear el proceso de aprendizaje de los estudiantes en las plataformas web, es evidente que aún es necesaria la implementación de analíticas de aprendizaje que involucren la mayor cantidad posible de interacciones del estudiante con las actividades y recursos de un curso, con el afán de realizar una adecuada monitorización del proceso de aprendizaje (Sánchez, García Saiz, & Zorrilla, 2012). En consecuencia, en este estudio se busca responder las siguientes interrogantes:

P.I.1 ¿Qué secuencias de aprendizaje se pueden descubrir a partir de las interacciones de los estudiantes en un curso virtual de Moodle?

P.I.2 ¿Qué variables son útiles para predecir el abandono de estudiantes en un curso virtual de Moodle?

P.I.3 ¿Cómo implementar un dashboard para visualizar el comportamiento de estudiantes en un curso virtual de Moodle?

1.4. Objetivos generales y específicos

Objetivo General

Evaluar un dashboard de análisis de comportamiento y predicción de abandono de los estudiantes en la plataforma Moodle.

Objetivos específicos



- Desarrollar un análisis exploratorio del comportamiento de los estudiantes utilizando técnicas de Machine Learning para descubrir secuencias de aprendizaje basado en las interacciones de los estudiantes con la plataforma.
- Identificar las variables más importantes para predecir el abandono de estudiantes (demográficas, notas, interacción, secuencias de aprendizaje).
- Implementar un dashboard en Moodle para visualizar el comportamiento de los estudiantes.
- Evaluar la usabilidad del dashboard con docentes en un contexto real.

1.5. Metodología

La metodología empleada en el estudio consta de seis etapas basadas en dos enfoques: primero, la extracción del comportamiento y segundo, la visualización de tal comportamiento.

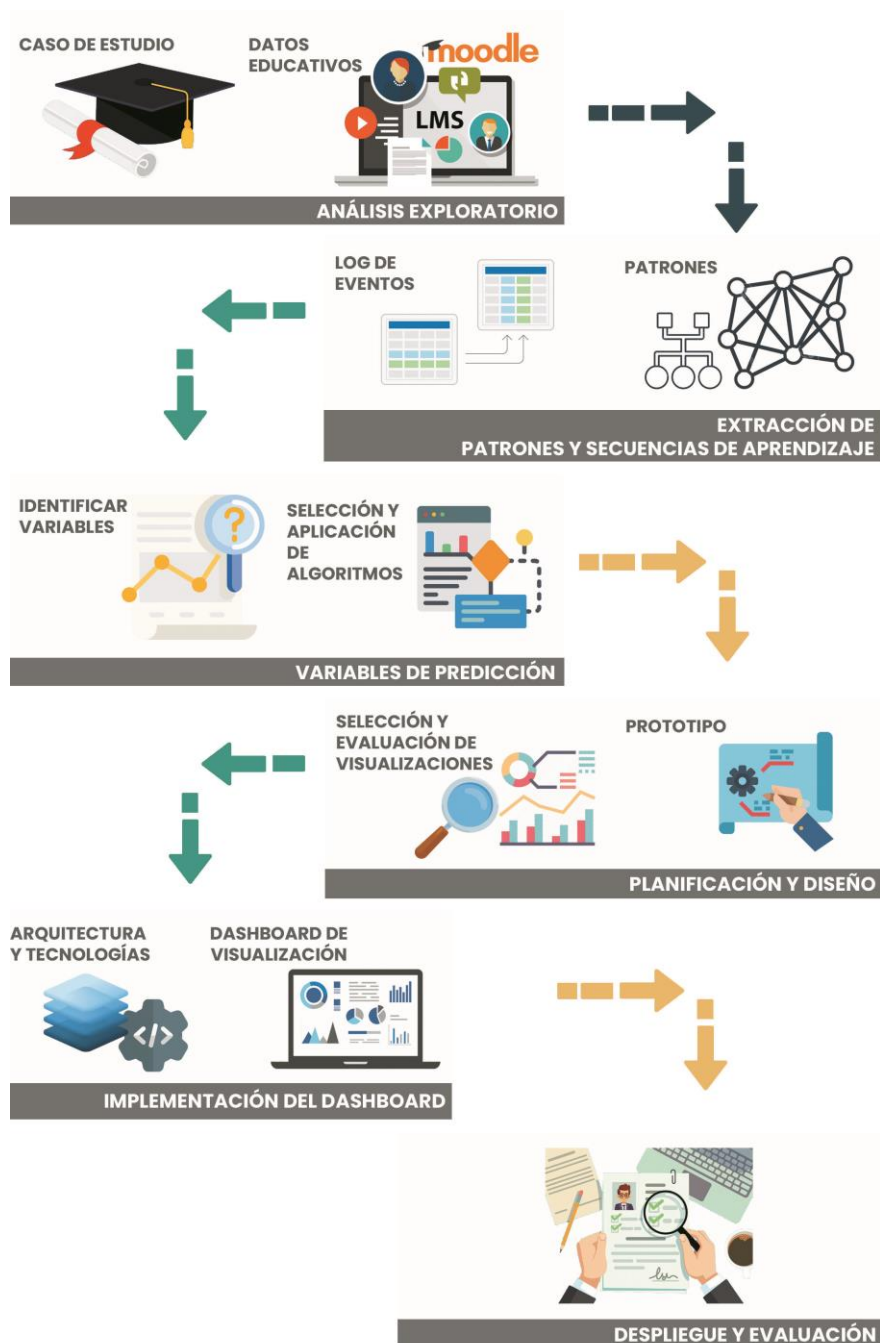
El primer enfoque consiste en la exploración de interacciones, hallazgo de patrones de comportamiento de los estudiantes y detección de variables de predicción de abandono en la plataforma Moodle. Por lo tanto, este enfoque comprende las etapas de Análisis Exploratorio de los Datos, Identificación de Patrones de Comportamiento y Detección de Variables de Predicción de Abandono. El segundo enfoque consiste en el diseño del dashboard, el proceso de implementación y la evaluación de la herramienta.

A continuación, se detallan las actividades a desarrollar dentro de estas etapas:

- Análisis Exploratorio de los Datos
 - Descripción del Caso de Estudio utilizado.
 - Análisis de Requerimientos de Diseño.
 - Extracción los datos educativos (estáticos y dinámicos) a través del análisis de las fuentes de datos.
- Identificación de Patrones de Comportamiento
 - Generación de un log de eventos de las interacciones de los estudiantes.
 - Extracción de patrones y secuencias de aprendizaje mediante Minería de Procesos.
 - Interpretación de los patrones y secuencias de aprendizaje encontradas.
- Detección de Variables de Predicción de Abandono
 - Identificación de variables de predicción en Moodle.
 - Selección de algoritmo de Minería de Datos para predecir el abandono estudiantil.



- Aplicación de Técnicas de Minería de Datos.
 - Definición de intervalos de tiempo para la predicción de abandono.
- Planificación y Diseño del Dashboard
 - Búsqueda y selección de visualizaciones existentes
 - Evaluación de visualizaciones
 - Prototipado de la Herramienta
- Implementación del Dashboard
 - Selección de la arquitectura de la herramienta.
 - Selección de las tecnologías adecuadas.
 - Implementación de la herramienta.
- Evaluación del Dashboard.
 - Selección de los actores involucrados en la evaluación.
 - Despliegue del dashboard a los actores seleccionados.
 - Aplicación de cuestionarios a los evaluadores sobre criterios de usabilidad.
 - Interpretación de los resultados de la evaluación.

Figura 1*Etapas de la Metodología empleada*

Nota. Marco metodológico propuesto para el desarrollo del dashboard de visualización.



1.6. Estructura de la Tesis

Esta tesis se encuentra estructurada de la siguiente manera:

El capítulo dos introduce el impacto de las tecnologías de información y comunicación en la educación, para esto el capítulo presenta los diferentes modelos de aprendizaje y los sistemas de gestión de aprendizaje denominados LMS (*Learning Management System – en inglés*).

El capítulo tres presenta las características del LMS Moodle, donde se describe sus particularidades, las actividades, recursos y bloques que forman parte de su estructura. Además, se presenta la arquitectura sobre la que se basa Moodle, el sistema de directorios, los datos y una breve descripción de la base de datos. Además, se presenta una descripción general de las tablas más importantes, las cuales servirán de insumo para generar el dashboard de visualización.

El capítulo cuatro aborda las analíticas de aprendizaje en Moodle. Se definen las analíticas de aprendizaje, se presentan herramientas o plugins existentes sobre Moodle, se presentan los patrones de aprendizaje sobre Moodle, además la predicción en base a interacciones y finalmente métodos de visualización como los dashboard.

El capítulo cinco presenta el análisis exploratorio de los datos donde se muestran los principales conceptos teóricos relacionados al análisis de datos para la construcción del dashboard de visualización y predicción de deserción estudiantil.

El capítulo seis aborda el proceso de desarrollo llevado a cabo para diseñar e implementar el dashboard de visualización para Moodle. Iniciando por la fase de planificación y diseño, la fase de implementación donde se describe la arquitectura implementada, tecnologías utilizadas, proceso de instalación/desinstalación, descripción de sus funcionalidades, características, estructura y finalmente la fase de evaluación.

Finalmente, el capítulo siete presenta las principales conclusiones de esta tesis y las líneas de trabajo futuro.

1.7. Contribuciones de la Tesis

A continuación, se detallan las principales contribuciones realizadas por este trabajo de Tesis:

(1) Escritura y publicación de un Paper de Posición de la propuesta del Trabajo de Investigación. Esto se realizó previo al desarrollo de la herramienta implementada. El artículo fue presentado de manera virtual en la XV Conferencia Latinoamericana de Tecnologías de Aprendizaje CLEI XLVI - LACLO XV (véase Anexo 1). Esta conferencia ocurre anualmente desde el 2006, el



mismo año cuando fue creada la comunidad Latinoamericana de Objetos de Aprendizaje (LACLO). Ha sido organizada en diferentes países de Latinoamérica, y en 2020 tuvo lugar en Ecuador.

(2) Publicación del artículo en el repositorio digital de IEEE ([Proposal for the Design and Evaluation of a Dashboard for the Analysis of Learner Behavior and Dropout Prediction in Moodle](#), véase Anexo 2).

(3) Diseño e implementación de una extensión (plugin) para Moodle que permite visualizar tanto al docente como al estudiante diferentes indicadores relacionados al proceso de aprendizaje. Además, el docente tiene la capacidad de analizar el comportamiento de los estudiantes de manera que pueda predecir el abandono estudiantil de manera temprana.



Capítulo 2: Tecnologías de la Información y Comunicación en la Educación

2.1. Aprendizaje Presencial

El proceso educativo tiene como elemento central al aprendizaje. Un modelo de aprendizaje es un sistema basado en una metodología compuesta de rasgos, estrategias y pautas propias que han sido diseñadas con el objetivo de orientar el proceso de aprendizaje (García Alba, 2010). Los modelos de aprendizaje han intentado responder a una necesidad histórica en el ámbito de la enseñanza: cómo potenciar el aprendizaje, cuáles son los factores de motivación o cuáles serían las características personales de cada estudiante. La mayor parte de ellos han sido aplicados al mundo de la enseñanza formal o se derivan directamente de la observación de cómo se aprende en dicho entorno (Westbrook, 2006).

El modelo pedagógico tradicional se caracteriza por la marcada diferencia de roles entre el alumno y el profesor (Maldonado Mahauad, Perez Sanagustín, Kizilcec, Morales, & Munoz Gama, 2017). En este tipo de sistema educativo el alumno es un receptor pasivo de la información, mientras que todo el peso del proceso educativo recae en el profesor. A pesar de su popularidad en tiempos pasados, el modelo pedagógico tradicional no está exento de críticas. Con el paso del tiempo, tanto estudiante como los propios cuerpos de docentes, reclaman que este ha quedado obsoleto; siendo considerado como un modelo predecible, poco estimulante y que necesita una adaptación urgente a los nuevos tiempos (Abeysekera & Dawson, 2014).

2.2. Aprendizaje Híbrido

El uso de entornos virtuales en la educación superior ha permitido implementar un enfoque de modalidades híbridas (Buenaño Fernández & Luján Mora, 2016). Este enfoque aumenta las posibilidades de crear un ambiente de aprendizaje efectivo y obtener respuestas más positivas de los estudiantes que el aprendizaje puramente virtual o e-learning (Westbrook, 2006). Por lo tanto, el aprendizaje híbrido, constituye una herramienta útil para distribuir la carga de trabajo de los estudiantes entre el aula y el hogar, de acuerdo con las directivas europeas para la educación superior.

El aprendizaje híbrido es considerado como un enfoque que combina los beneficios ofrecidos por los componentes de aprendizaje presencial y en línea (Abubakar, Kamsin, & Abdullah, 2020). El proceso de aprendizaje híbrido se enriquece con la interacción y la discusión que se produce entre los materiales docente-estudiante, estudiante-estudiante y estudiante-aprendizaje (Buenaño Fernández &



Luján Mora, 2016). La razón principal del interés en el diseño de entornos de aprendizaje híbrido es que se ha encontrado que las combinaciones de actividades de enseñanza presencial y en línea ofrecen varias oportunidades nuevas para optimizar el aprendizaje (Spanjers, y otros, 2015). En particular, el aprendizaje híbrido respaldado por plataformas virtuales se está convirtiendo en una de las perspectivas de aprendizaje y enseñanza más impactantes en la educación superior (Gutiérrez Braojos, Montejó Gamez, Marin Jimenez, & Campaña , 2018).

2.3. Aprendizaje en Línea

El rápido crecimiento y mejora en las tecnologías en línea en los últimos años ha provocado la aparición y desarrollo del aprendizaje virtual (Gutiérrez Braojos, Montejó Gamez, Marin Jimenez, & Campaña , 2018). Actualmente el aprendizaje en línea es una opción de gran acogimiento por parte de numerosas instituciones presenciales o a distancia (G. de Barba, y otros, 2020). La característica principal de este aprendizaje es que las interacciones se dan en un espacio virtual a través del uso de las tecnologías de la información y la comunicación (Gazzotti, y otros, 2011). Esta modalidad de enseñanza-aprendizaje consiste en el diseño, puesta en práctica y evaluación de un curso o plan formativo (Rabbany, Takaffoli, & Zaïane, 2011). Comúnmente se denomina Aprendizaje Electrónico (*e-learning – Electronic Learning en inglés*). El e-learning se está expandiendo rápidamente por todo el sistema educativo impregnando la educación formal y no formal, así como la enseñanza presencial y la formación a distancia (Llorente Cejudo, 2007).

2.4. Sistemas de Gestión de Aprendizaje

Las plataformas de aprendizaje virtual hacen parte de una tendencia de la educación, presente desde mediados de los años noventa (Gutiérrez Braojos, Montejó Gamez, Marin Jimenez, & Campaña , 2018). Estas plataformas se convierten en herramientas tecnológicas de apoyo a la enseñanza virtual y permiten compartir recursos didácticos estructurados en cursos en línea, acompañando el proceso tradicional de enseñanza-aprendizaje-evaluación (Bogarín, Romero, Cerezo, & Sánchez Santillán, 2014).

Con el desarrollo y la creciente popularidad de los entornos de aprendizaje compatibles con la tecnología, los sistemas de información permiten capturar todos los eventos/acciones/actividades de los estudiantes en diferentes niveles de granularidad (Valle Torre, Tan, & Hauff, 2020), desde eventos de bajo nivel como pulsaciones de teclas, gestos del mouse y clics, hasta niveles superiores eventos tales como las actividades de aprendizaje de los estudiantes (Bogarín, Cerezo, & Romero, A survey on educational process mining, 2017). Estos entornos de aprendizaje acumulan una gran

cantidad de información que es muy valiosa para analizar la interacción y comportamiento de los estudiantes (Mostow & Beck, 2006).

En los últimos años, los sistemas informáticos integrados conocidos como LMS han surgido rápidamente y están teniendo, y tendrán cada vez más, profundos efectos en la enseñanza y el aprendizaje universitarios (Bogarín, Cerezo, & Romero, A survey on educational process mining, 2017). El uso de estos sistemas educativos ha crecido exponencialmente, estimulado por el hecho de que ni los estudiantes ni los docentes están obligados a una ubicación específica (Yassine, Kadry, & Sicilia, 2016). Dichos sistemas de aprendizaje electrónico a veces también se conocen como CMS, LCMS, Entorno de Aprendizaje Gestionado (*MLE – Managed Learning Environment en inglés*), Sistema de Apoyo al Aprendizaje (*LSS – Learning Support System en inglés*) o Plataforma de Aprendizaje (*LP – Learning platform en inglés*) (Romero, Ventura, & Garcia, 2008). Como resultado, los Entornos de Aprendizaje Virtual (*VLE – Virtual Learning Environment en inglés*) se instalan cada vez más en universidades, colegios, escuelas, empresas e incluso instructores individuales para agregar tecnología web a sus cursos y complementar los cursos tradicionales presenciales (García Alba, 2010).

A pesar de las capacidades que las plataformas LMS aportan al proceso del aprendizaje, éstas se encuentran infravaloradas (Costa, Alvelos, & Teixeira, 2012). La subutilización de las plataformas LMS se debe en gran medida al hecho de que estos sistemas únicamente almacenan registros de interacción, aunque con el surgimiento de áreas de investigación como la analítica de aprendizaje y minería de procesos educativos se les ha dado más atención en busca de explorar el potencial de los datos almacenados por los entornos de aprendizaje LMS (Juhaňák, Zounek, & Rohlíková, 2019). Una razón que justifica el bajo uso de las plataformas LMS por parte de los docentes es la falta de herramientas que permitan conocer y analizar las interacciones del estudiante en la plataforma virtual, por lo que sería útil el desarrollo de una herramienta que les permita visualizar el comportamiento de los estudiantes en la plataforma y que contribuya a la calificación del alumno (Juhaňák, Zounek, & Rohlíková, 2019). De la misma manera, podrían utilizarse los resultados del análisis de los datos almacenados por las plataformas LMS para el rediseñar cursos existentes (Den Beemt, Buijs, & Der Aalst, 2018), mejorar los procesos de aprendizaje y predecir la deserción estudiantil (Moreno Marcos, Muñoz Merino, Alario Hoyos, Estévez Ayres, & Delgado Kloos, 2018).



Capítulo 3: Sistema de Gestión de Aprendizaje Moodle

3.1. Moodle como EVEA

El aprendizaje en línea, incluidos los cursos combinados y totalmente en línea, se ha convertido en un aspecto común de la educación en las últimas dos décadas (Al-Ajlan & Zedan, 2008). Este aprendizaje en línea ha traído consigo desafíos sobre cómo mejorar el proceso de enseñanza y aprendizaje (Dreizzen, Zangara, & Alonso, 2016). Como resultado del crecimiento generalizado y la popularidad del aprendizaje en línea, las herramientas denominadas Entornos Virtuales de Enseñanza y Aprendizaje (*EVEA*) promueven la comunicación pedagógica entre los actores del proceso educativo, sea un aprendizaje presencial, en línea o combinado (Dreizzen, Zangara, & Alonso, 2016). El concepto de EVEA, aparece como evolución del concepto de Plataforma de Tele formación, el cual denomina precisamente los entornos diseñados para automatizar y gestionar el desarrollo de actividades formativas, también conocidos como LMS (Gazzotti, y otros, 2011). Los EVEAS integran diferentes perfiles, en función de las necesidades y requerimientos de distintos públicos que son usuarios de estas aplicaciones: docentes, estudiantes y administradores (Llorente Cejudo, 2007). Pueden soportar diferentes tareas asociadas a la enseñanza y el aprendizaje, como por ejemplo la gestión y distribución de contenidos, la comunicación sincrónica y asincrónica, repositorios de materiales individuales y grupales, trabajo colaborativo, evaluación y la gestión y seguimiento de la enseñanza (Castañeda Quintero & López Vicent, 2007).

En este contexto, Moodle es una plataforma especializada en contenidos de aprendizaje y uno de los sistemas de gestión de cursos y contenido más populares (García Alba, 2010). Para (Al-Ajlan & Zedan, 2008), Moodle es una herramienta polivalente que ofrece un amplio espectro de comunicación didáctica. Este LMS es distribuido bajo la Licencia Pública General de GNU, es decir, si tiene derechos de autor, pero viene con autorización para que cualquiera pueda usarlo, copiarlo y distribuirlo (Einhardt, Aires Tavares, & Cechinel, 2016). Fue creado por Martin Dougiamas para ayudar a los educadores a crear cursos en línea con un enfoque en la interacción y la construcción colaborativa de contenido (Luna, Castro, & Romero, 2017). Ofrece varias herramientas de evaluación que miden diferentes aspectos en el proceso de aprendizaje, como registros de actividades, participación en cursos, cuestionarios y evaluaciones, análisis de compromiso y muchos más (Moodle, 2020). Cuenta con soporte para más de 100 idiomas y contiene prestaciones que se consideran necesarias en el marco de la educación superior: foros, agendas, chat, wikis, evaluaciones estructuradas, difusión de recursos, entre otros. Moodle ha sido estudiada por diversos organismos a nivel mundial que la han puesto en un lugar privilegiado dentro de las opciones de EVEAS.



disponibles y además cuenta con una comunidad virtual que la soporta y la hace crecer de manera cooperativa (Barrios, Fernández, Godoy, & Mariño, 2012). Además, permite implementar Internet en la docencia presencial, híbrida y virtual, destacando la facilidad en el manejo, su estabilidad y versatilidad.

En la actualidad, el uso de los EVEAS como Moodle en complemento de la enseñanza presencial ha dado origen a diferentes tipos de manifestaciones educativas, que extienden las posibilidades del aula presencial, incorporando opciones de enseñanza mediadas por las herramientas ofrecidas en estos entornos (Dreizen, Zangara, & Alonso, 2016). Uno de estos escenarios de aprendizaje es la denominada propuesta pedagógica de Aula Invertida (*FC - Flipped classroom en inglés*), donde el aprendizaje requiere la participación activa de los estudiantes en las actividades de aprendizaje antes, durante y después de las sesiones de estudio (Lage, Platt, & Treglia, 2000). Sin embargo, los estudiantes con frecuencia carecen de las habilidades, el tiempo y/o la motivación necesarios para participar plenamente en las actividades previas a la clase y, por lo tanto, no se comprometen con el nivel de participación en el proceso de aprendizaje que complementa efectivamente el diseño previsto. Según (Cockrum, 2014) el uso del modelo pedagógico de FC apoyado con tecnología permite hibridar el espacio de aprendizaje y llegar de forma personalizada a los estudiantes, adaptándose a su ritmo de aprendizaje y cubriendo sus demandas al momento de aprender los contenidos.

3.2. Características

Moodle, como su página web (www.moodle.org) indica, es un paquete de software para la creación de cursos y sitios web basados en Internet (Moodle, 2020). Es una herramienta inspirada en la pedagogía social-constructivista, la cual plantea la idea de que el conocimiento se va construyendo por el estudiante a partir de su participación activa en el proceso y en relación con su entorno social (Valenzuela Zambrano & Pérez Villalobos, 2013). Moodle posee las siguientes características (Moodle, 2020):

- *Mundialmente probado y de confianza:* Moodle ha impulsado cientos de miles de ambientes de aprendizaje globalmente, Moodle tiene la confianza de instituciones y organizaciones grandes y pequeñas. El número de usuarios de Moodle a nivel mundial, de más de 200 millones de usuarios registrados en más de 212 países (en agosto del 2020), entre usuarios académicos y empresariales, lo convierten en la plataforma de aprendizaje más ampliamente utilizada del mundo (más del 70% de las instituciones de educación superior según estudios).



- Diseñado para soportar tanto la enseñanza como el aprendizaje: Con más de diez años de desarrollo guiado por la pedagogía de constructivismo social, Moodle proporciona un conjunto poderoso de herramientas centradas en el estudiante y ambientes de aprendizaje colaborativo, que le dan poder, tanto a la enseñanza como al aprendizaje.
- *Gratuito, sin cargos por licenciamiento:* Moodle es proporcionado gratuitamente como programa de código abierto. Cualquier persona puede adaptar, extender o modificar Moodle, tanto para proyectos comerciales como no-comerciales, sin pago de cuotas por licenciamiento.
- *Plataforma de aprendizaje todo-en-uno:* Moodle presenta un amplio abanico de diseño de actividades pedagógicas para quienes quieran realizar aprendizaje completamente en línea como también semi presencial (b-learning).
- *Altamente flexible y personalizable:* Debido a que es código abierto, Moodle puede ser personalizado en cualquier forma deseada, para adecuarlo a necesidades individuales. Su configuración modular y diseño interoperable permite a desarrolladores crear plugins e integrar aplicaciones externas para lograr funcionalidades específicas.
- *Escalable a cualquier tamaño:* Desde unos cuantos estudiantes hasta millones de usuarios, Moodle puede escalarse para soportar las necesidades, tanto de clases pequeñas, como de grandes organizaciones. Debido a su flexibilidad y escalabilidad, Moodle ha sido adoptado para usarse en educación, negocios, organizaciones no-lucrativas y contextos comunitarios.
- *Perfiles de usuario:* El funcionamiento de Moodle se basa en la interacción de cuatro tipos de usuarios: invitados, estudiantes, docentes y administradores. Los invitados son autorizados por el administrador(es) y por el docente(s), es el grupo que menos privilegios tiene, por ende, su accionar es limitado. Los estudiantes, en cambio, pueden matricularse en los cursos, participar en las actividades y utilizar sus recursos, así como también formar grupos para interactuar entre ellos y con el profesor. Los administradores poseen todos los privilegios y su principal función es gestionar la información de la base de datos y controlar su acceso. Finalmente, los profesores son los que diseñan las actividades y los materiales de las asignaturas, con base en la aplicación de principios pedagógicos. Moodle les permite controlar y evaluar el aprendizaje de cada estudiante y realizar seguimiento de sus avances (Valenzuela Zambrano & Pérez Villalobos, 2013).
- *Orientado a objetos de aprendizaje:* Estos objetos de aprendizaje están diseñados para distribuirse a través de internet posibilitando el acceso simultáneo a la información por parte de múltiples usuarios.



3.2.1. Actividades

Una actividad es un nombre general para un grupo de características en un curso Moodle. Usualmente una actividad es algo que un estudiante hará y recibirá una calificación por ello. Estas actividades permiten la interacción con otros estudiantes o el docente. En la terminología de Moodle, una Actividad, como, por ejemplo, Foro o Examen, significa propiamente algo a lo que los estudiantes pueden contribuir directamente, y a menudo es contrastada con un recurso, como por ejemplo un archivo o una página, el cual es presentado por el docente a los alumnos (Moodle, 2020). Sin embargo, el término actividad en ocasiones por conveniencia también es usado para referirse tanto a actividades como recursos. Hay 14 diferentes tipos de actividades en Moodle 2.x y 3.x estándar (Moodle, 2020) como se puede observar en la Tabla 1. Adicionalmente pueden instalarse complementos (plugins de terceros) desde la base de datos del subdirectorío de Plugins de Moodle para añadirlos al sitio (Castañeda Quintero & López Vicent, 2007).

Tabla 1

Descripción de las Actividades Disponibles en Moodle

Actividad	Descripción
Tareas	Tipo de actividad que proporciona un espacio en el que los estudiantes pueden enviar sus trabajos para que los docentes los califiquen y proporcionen retroalimentación.
Chat	El chat de Moodle permite mantener conversaciones entre usuarios en tiempo real. Incluye fotos de los perfiles en la ventana de chat y soporta direcciones URL, emoticonos, integración de HTML, imágenes, entre otras.
Elección	Tipo de actividad en la que un docente hace una pregunta y especifica una variedad de respuestas de opción múltiple relacionadas con los conceptos cubiertos en algún tema del curso. Los estudiantes pueden responder cada pregunta, revisar si la respuesta fue correcta o solicitar ver todas las respuestas.
Base de Datos	El módulo Base de Datos permite al docente y los estudiantes construir un banco de registros sobre cualquier tema o asunto, así como realizar búsquedas y mostrar resultados. La estructura de las entradas de esta base de datos incluye imágenes, archivos, direcciones URL, números, texto entre otras cosas.
Retroalimentación	Actividad que le permite a los docentes crear y aplicar encuestas a los estudiantes, con el propósito de conocer su opinión sobre un tema.
Foro	Tipo de actividad que permite a los participantes tener discusiones asíncronas entre ellos de un tema específico planteado por el docente.
Glosario	El glosario presenta un catálogo de palabras relativas a un campo de estudio, actividad, tema o área del saber, con su definición y/o comentarios.
Lección	Una lección proporciona contenidos interactivos en una serie de páginas (HTML integrado en Moodle), y textos que el alumno debe recorrer. Al final de cada página se puede plantear una pregunta con varias posibles respuestas para comprobar de alguna manera que el alumno lo ha leído y comprendido.



(LTI) Herramienta Externa	La herramienta externa les permite a los participantes interactuar con actividades y recursos de aprendizaje compatibles con LTI (<i>LTI – Learning Tools Interoperability en inglés</i>) en otros sitios web. Una herramienta externa podría proporcionar acceso a un nuevo tipo de actividad, o material de aprendizaje de algún editor.
SCORM	Un paquete SCORM (<i>SCORM – Sharable Content Object Reference Model en inglés</i>) es un bloque de material web empaquetado de una manera que sigue el estándar SCORM de objetos de aprendizaje. Estos paquetes pueden incluir páginas web, gráficos, programas JavaScript, presentaciones Flash y cualquier otra cosa que funcione en un navegador web. De esta manera, es fácil agregar cualquier paquete SCORM como parte del curso.
Encuesta Predefinida	Proporciona instrumentos de encuesta útiles en la evaluación y estimulación del aprendizaje en ambientes en línea. Dos de ellos son Encuesta Constructivista sobre el Entorno de Aprendizaje en Línea (<i>COLLES – Constructivist On-Line Learning Environment Survey en inglés</i>) y Encuesta de Actitudes hacia el Pensamiento y el Aprendizaje (<i>ATTLS – Attitudes to Thinking and Learning Survey en inglés</i>).
Wiki	Un Wiki posibilita la creación colectiva de documentos en un lenguaje simple de marcas, utilizando un navegador web.
Taller	Actividad en la cual los estudiantes envían su propio trabajo y reciben varios envíos de otros estudiantes, los que tienen que evaluar de acuerdo con las especificaciones del docente.
Examen	Tipo de actividad que permite al docente diseñar y armar exámenes, que pueden ser calificados automáticamente o se puede dar retroalimentación o mostrar las respuestas correctas.

3.2.2. Recursos

Los recursos en Moodle pueden definirse como aquellos contenidos que no suponen una entrega calificable por parte del estudiante (Valenzuela Zambrano & Pérez Villalobos, 2013). Simplemente son materiales que el profesor pone a disposición del alumnado como apuntes de la asignatura (todo tipo de documentos), o enlaces a recursos externos. Moodle divide sus recursos en tres grupos: gestión de contenidos, comunicación y evaluación (Ros Martínez, 2008). Los recursos que ofrece Moodle (Moodle, 2020) pueden observarse en la Tabla 2.

Tabla 2

Descripción de los Recursos Disponibles en Moodle

Recurso	Descripción
Archivo	El tipo de recurso archivo puede ser una imagen, un documento PDF, una hoja de cálculo, un archivo de sonido, un archivo de video.
Carpeta	El tipo de recurso carpeta ayuda a los estudiantes y docentes a organizar los archivos. Las carpetas pueden contener otras carpetas.



Etiqueta	El tipo de recurso etiqueta pueden ser palabras o una imagen para separar recursos y actividades en un tema o una lección, aunque también pueden ser descripciones largas o instrucciones para las actividades.
Libro	Este tipo de recurso multi página tiene un aspecto similar a un libro. Los docentes pueden exportar sus Libros como paquete IMS (tipo de formato de archivo estándar basado en una serie de especificaciones que facilitan la reutilización de contenidos en distintos sistemas sin necesidad de convertirlos a otro formato).
Página	Este recurso permite al estudiante observar una página navegable y simple que el docente crea con un editor HTML.
Paquete de contenido IMS	Este tipo de recurso agrega material estadístico desde otros recursos en el formato IMS.
URL	Este tipo de recurso permite a los estudiantes acceder a cualquier sitio web a través del navegador.

3.2.3. Bloques

La plataforma de Moodle cuenta con bloques que le permiten al estudiante obtener información rápida del curso y acceder a algunas herramientas útiles; por ejemplo, puede ver las novedades, calificaciones, eventos del calendario, participantes del curso, contador de visitas, búsquedas en foros, calculadora, comentarios rápidos, entre otros (Castañeda Quintero & López Vicent, 2007). Por defecto, la plataforma lleva instalados varios, pero existen muchos más en la página oficial de Moodle que pueden ser descargados e instalarlos (Ros Martínez, 2008). Hay diferentes tipos de bloques que pueden ser usados tanto por el docente como por el estudiante (Moodle, 2020) como se puede ver en la Tabla 3 (no se incluyen todos los disponibles, sino aquellos más comunes o usuales).

Tabla 3

Descripción de los Distintos Tipos de Bloque en Moodle

Bloque	Descripción
Bloque de Actividades	El bloque de actividades lista todas las categorías de las actividades disponibles en el curso (foros, tareas, etc.). Con el primer acceso al curso únicamente se mostrarán los foros, debido a que únicamente se ha creado el foro de Novedades. Otros tipos de actividades se muestran conforme se van añadiendo al curso.
Bloque de Administración	Los enlaces en este bloque únicamente están disponibles para el docente. Este bloque permite a los docentes gestionar las matrículas de los alumnos en el curso, ver los grupos conformados, ver un libro de calificaciones, crear escalas de calificación y tener acceso al foro del docente.
Bloque de Calendario	Este bloque presenta los eventos en el curso. Todas las actividades en Moodle con una fecha límite serán eventos en el calendario de forma automática. Los eventos pueden implicar usuarios individuales, grupos definidos o cursos completos.



Bloque de Cursos	Este bloque muestra un listado de todos los cursos en los que un estudiante se encuentra matriculado o cursos asignados a un docente y permite acceder a la página principal de otro curso.
Bloque HTML	Este bloque permite agregar texto y gráficos. Es un bloque flexible que puede incorporar variedad de funciones y usos. Proporciona un editor HTML para formatear texto, agregar imágenes, crear enlaces, agregar videos, sonido y otros ficheros.
Bloque de Mensajes	Este bloque constituye un sistema de mensajería interna de Moodle. Permite intercambiar mensajes entre los usuarios.
Bloque de Usuarios en Línea	Este bloque muestra los usuarios que se han registrado en el curso actual por un periodo de tiempo fijado por el administrador del curso (por defecto cinco minutos).
Bloque de Personas y Participantes	Este bloque contiene un enlace a participantes que es un listado de todos los usuarios del curso.
Bloque de Búsqueda en Foros	Este bloque cuenta con un buscador que permite encontrar una palabra o una frase en los foros del curso.
Bloque de Eventos Próximos	Este bloque muestra una lista de los acontecimientos próximos en el calendario, con enlaces al contenido del acontecimiento señalado. Los eventos son generados de manera automática por el calendario o por la fecha de finalización de las actividades.
Bloque de Canales RSS Remotos	Estos canales RSS poseen un formato que cumple con el estándar XML para compartir contenido en la Web, se usan para difundir información a usuarios que se han suscrito a una fuente de contenidos.

3.3. Arquitectura

Moodle ha sido creado con una infraestructura de internet LAMP que consiste en Linux (sistema operativo), Apache (servidor web), MySQL (base de datos) y PHP (lenguaje de programación) (Wild, 2017). Si bien Moodle puede ejecutarse en otras pilas de tecnología, LAMP ha demostrado ser la configuración más popular entre los administradores de Moodle (Moore & Churchward, 2010). Debido a la portabilidad de estos componentes y la modularidad de Moodle en sí, puede admitir una amplia gama de sistemas operativos, sistemas de bases de datos y servidores web (Rice, 2015). La Figura 2 muestra una descripción general de la arquitectura de Moodle.

El nivel más bajo es el sistema operativo. Si bien Linux es la plataforma preferida, se admiten otros derivados de Unix como Solaris y AIX, junto con Windows y Mac OS X (preferiblemente las variantes de servidor para sitios de producción) (Büchner, 2016). PHP es el lenguaje de programación en el que está escrito Moodle (acompañado de archivos HTML, JavaScript y CSS). Es el único componente que no se puede reemplazar con ninguna otra contraparte (Wild, 2017). MySQL es la base de datos elegida para la mayoría de las aplicaciones de código abierto, pero otros sistemas de bases de datos como Microsoft SQL Server, Oracle, PostgreSQL y MongoDB son totalmente compatibles (Büchner, 2016). Apache se ha convertido en el estándar para aplicaciones web a gran

escala, seguido de cerca por Microsoft IIS. Ambos servidores web son compatibles como cualquier otro que ofrezca compatibilidad con PHP (Rice, 2015). La interacción de los elementos de la arquitectura de Moodle se muestra en la Figura 3.

Figura 2

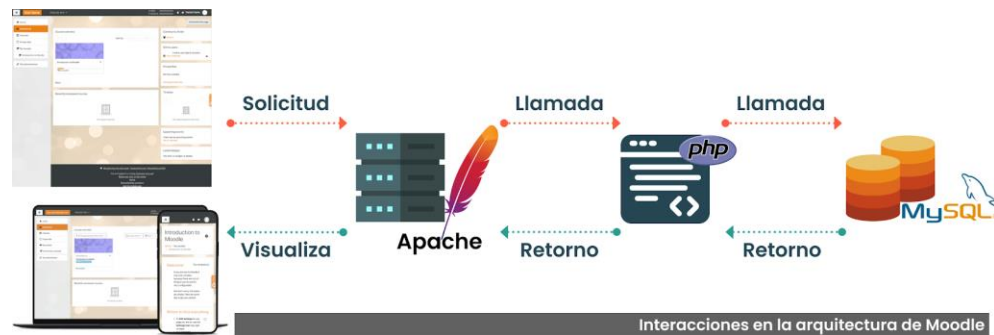
Descripción General de la Arquitectura de Moodle



Nota. Adaptada de (Wild, 2017).

Figura 3

Interacción de los Elementos de la Arquitectura de Moodle



Nota. Adaptada de (Büchner, 2016).

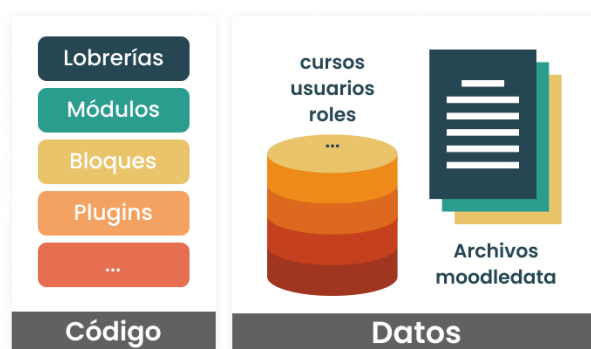
El usuario realiza solicitudes a través de la interfaz del navegador web o una aplicación móvil de Moodle (por ejemplo, para mostrar un recurso de aprendizaje). El navegador web pasa la solicitud al servidor web, que llama al módulo PHP responsable de la llamada. El módulo PHP llama a la base de datos con una acción (operación de consulta, actualización, inserción o eliminación) que devuelve los datos solicitados. Con base en esta información, el módulo PHP devuelve datos (generalmente en

forma de código HTML o JavaScript) a los servidores web, los cuales pasan la información que se mostrará al navegador o aplicación del usuario (Büchner, 2016).

En la capa de Moodle se hacen distinciones entre código (principalmente escrito en PHP, HTML y CSS) y datos (principalmente agregado a través de la interfaz de Moodle) como se muestra en la Figura 4.

Figura 4

Capa de Moodle de la Arquitectura General



Nota. Adaptada de (Büchner, 2016).

Las bibliotecas, módulos (como recursos y actividades), bloques, complementos y otras entidades de Moodle se representan en código. Siempre se almacena en el sistema de archivos en un directorio de Moodle denominado `dirroot`, que se especifica durante el proceso de instalación (Büchner, 2016). El código incluye todos los elementos que se ocupan de las operaciones de backend (servidor) y frontend (interfaz de usuario). Los cursos, usuarios, roles, grupos, calificaciones y otros datos de Moodle, como los recursos de aprendizaje agregados por los profesores, las publicaciones en foros agregados por los estudiantes y las configuraciones del sistema agregadas por el administrador, se almacenan principalmente en la base de datos de Moodle (Wild, 2017).

Sin embargo, archivos como imágenes de usuario o tareas (`assign`) cargadas se almacenan en otro directorio de Moodle, conocido como `moodledata`, que se encuentra en un directorio llamado `dataroot`. La información sobre los archivos (metadatos como el nombre, la ubicación, la última modificación, la licencia y el tamaño) se almacena en la base de datos, que hace referencia a los archivos respectivos (Wild, 2017). Moodle gestiona sus archivos internamente mediante el mecanismo hash SHA1 (función hash criptográfica ampliamente utilizada, que genera 160 bits a

partir de cualquier valor de entrada), es importante enfatizar que la interferencia con cualquier archivo en moodledata interrumpirá la correcta ejecución del sistema (Büchner, 2016).

3.4. Estructura de Directorios y Sistema

Como se muestra en la Figura 5, cada sistema Moodle en funcionamiento se puede dividir en tres áreas separadas: el código de Moodle, la base de datos de Moodle y los datos de Moodle.

Figura 5

Áreas de Moodle



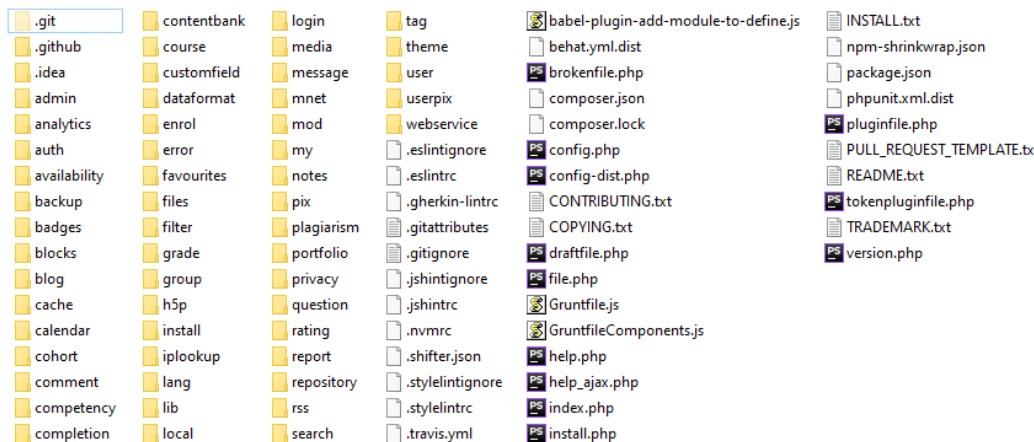
Nota. Adaptada de (Moore & Churchward, 2010).

3.4.1. Código de Moodle

Debido a que PHP es un lenguaje interpretado, el código de Moodle se almacena como archivos de código fuente en el servidor web (Moodle, 2020). Cuando se solicita un archivo en particular en el servidor, el intérprete de PHP analiza el código sobre la marcha y la salida resultante se envía a través del software del servidor web (Moore & Churchward, 2010). Cada carpeta de nivel superior representa un componente principal de Moodle. En algunos casos, los módulos también tendrán soporte para complementos adicionales. Un ejemplo de esto es el módulo de actividades de prueba, que admite tipos de preguntas modulares. Desde el punto de vista del usuario final, los módulos se instalan copiando el módulo en la ubicación de carpeta adecuada en el servidor (Moodle, 2020). Moodle detecta el nuevo módulo la próxima vez que un administrador ingresa al sistema, localiza el código SQL del módulo, lo ejecuta y finalmente muestra los resultados (Moore & Churchward, 2010). Las actualizaciones funcionan de la misma manera con Moodle, rastreando la versión de la base de datos y actualizando automáticamente la base de datos según sea necesario. Toda esta sencilla interfaz para los usuarios finales tiene el costo de un poco de esfuerzo para el desarrollador. La Figura 6 muestra la lista de directorios de una instalación reciente de Moodle 3.9.

Figura 6

Lista de Directorios de Moodle 3.9



Nota. Obtenida del directorio de instalación de Moodle. Elaboración Propia.

No es necesario mencionar todos los directorios de la carpeta principal, únicamente se explora las funciones de carpetas importantes utilizadas por desarrolladores (ver Tabla 4) que realizan modificaciones a Moodle. Moodle usa una nomenclatura simple para los módulos, donde cada uno están encerrado en su propia carpeta, y el nombre de la carpeta es el nombre que Moodle muestra en su interfaz cuando se refiere al módulo (Moodle, 2020).

Tabla 4

Descripción de los Módulos Principales en Moodle

Módulo	Descripción
admin	Esta carpeta almacena los archivos PHP que controlan la interfaz del usuario administrativo. También contienen el archivo <code>cron.php</code> , que se ejecuta como un proceso por lotes para realizar tareas de mantenimiento del sistema, como la entrega de mensajes y las copias de seguridad del curso.
auth	Esta carpeta contiene los módulos de autenticación de Moodle. Los módulos de autenticación controlan la creación de usuarios, los datos del perfil de usuario y el acceso de los usuarios al sistema.
backup	Esta carpeta contiene funciones de respaldo del curso para el sistema. No se trata de funciones de copia de seguridad para todo el sistema, sino de funciones para la copia de seguridad, restauración e importación de cursos. Cada módulo es autónomo, lo que nos permite, como desarrolladores, agregar módulos sin tener que modificar el código central.



blocks	Este módulo almacena la información de los bloques de Moodle . Este es uno de los tipos de módulos más simples de hacer y también tiende a funcionar en múltiples versiones de Moodle con pocas modificaciones o ninguna.
course	Este componente es de los más importantes dado que Moodle está organizado en torno a cursos. En este módulo los desarrolladores pueden modificar o agregar formatos e informes de cursos, formatos de curso y nuevos diseños a un curso.
enrol	Este directorio contiene todos los módulos de inscripción para Moodle. Estos módulos controlan la creación y gestión de asignaciones de roles a nivel de curso (inscripciones).
files	Este módulo permite incorporar archivos al sistema. Esto incluye la carga de archivos, el control de acceso y la visualización.
filter	El sistema de filtrado de Moodle es una función de búsqueda y reemplazo basada en texto / expresiones regulares. Los filtros hacen coincidir y modifican la página antes de que se muestre. Esta es una capacidad muy poderosa. Sin embargo, debe desarrollarse cuidadosamente, teniendo en cuenta las implicaciones de rendimiento.
lang	Este directorio almacena las cadenas de idioma del sistema principal. Esta es la base de la localización y el soporte de idiomas de Moodle. Todas las cadenas que se muestran al usuario final se asignan a través de esta función. Las asignaciones de cadenas de idioma también se almacenan en la carpeta de idioma de datos de Moodle. Esta estructura permite una fácil personalización local de los archivos de idioma.
lib	Este directorio almacena las funciones principales de la biblioteca del sistema. Con el desarrollo de módulos y personalizaciones, se usarán clases y funciones definidas en este directorio.
mod	Este directorio almacena módulos de actividades como tareas, cuestionarios, wiki, foros y módulos de lecciones. Las actividades de aprendizaje son el núcleo de cualquier curso impartido con Moodle.
my	Este directorio proporciona una lista de los cursos asignados a un alumno, incluido un resumen de las próximas actividades del curso. El usuario puede agregar y eliminar bloques en su página de portal. my proporciona una buena ubicación para mostrar información personalizada con cambios mínimos en el núcleo de Moodle.
theme	Este directorio almacena los temas integrados de Moodle y cualquier tema personalizado instalado en el sistema. Los temas son una combinación de CSS, HTML y PHP. Cada tema tiene su propia carpeta. Sin embargo, está limitado en cuanto a la parte de la página de Moodle que puede modificar ya que ciertos componentes de Moodle están codificados para mostrarse de cierta manera.

Tomado de (Moore & Churchward, 2010).

3.4.2. *Datos de Moodle*

Los datos de Moodle son la ubicación de almacenamiento de archivos para el contenido cargado por el usuario (Moodle, 2020). Moodle también almacena los datos de la sesión para los usuarios que han iniciado sesión en el sistema, si se utilizan sesiones basadas en archivos (Moore & Churchward, 2010). Los datos de Moodle también almacenan paquetes de idiomas opcionales que se pueden descargar desde la interfaz de administración de Moodle (Moodle, 2020). Moodle estructura



los datos en esta carpeta ya sea por el usuario o por el curso. La mayoría de los datos por tamaño y recuento de archivos están en los cursos. Cada curso tiene una carpeta, que se nombra con un valor entero (Moore & Churchward, 2010). El valor entero se establece en el `id` de la base de datos interna del curso en cuestión. Se puede determinar fácilmente estos valores navegando a un curso a través de la interfaz web de Moodle e inspeccionando su URL, por ejemplo:

<http://localhost/workspace/moodle39/course/view.php?id=3>

Dentro de una carpeta del curso, Moodle almacenará los datos del módulo en la carpeta `moddata` (Rice, 2015). Cuando un módulo necesita almacenar archivos, los guarda aquí en una carpeta con el mismo nombre que el módulo (Büchner, 2016). Por ejemplo, el módulo `wiki` tendrá una carpeta aquí llamada `wiki`. Además, Moodle creará una carpeta llamada `backupdata` si se ha creado alguna copia de seguridad del curso (Moodle, 2020). Cualquier archivo que haya sido cargado directamente por un usuario usando la interfaz de archivos del curso se cargará en la raíz de esta carpeta. Los usuarios también pueden crear sus propias carpetas y subcarpetas dentro de la carpeta `root` (Moodle, 2020).

Moodle 3.x utiliza un modelo organizativo completamente nuevo para archivos cargados por usuarios, que se basa en un algoritmo hash. El objetivo principal de este nuevo método es respaldar el uso eficiente del espacio de almacenamiento en disco y una mayor flexibilidad para usar archivos en múltiples cursos (Wild, 2017).

3.4.3. Base de Datos de Moodle

Moodle organiza sus tablas de datos en diferentes categorías con un propósito distinto dependiendo de su función específica en el sistema (Figura 7). En total, el modelo comprende 461 tablas interrelacionadas en su versión 3.9 (Moodle, 2020). Por ejemplo, la tabla `user` almacena los nombres, contraseñas y otra información sobre cada usuario de Moodle. La tabla `wiki_pages` almacena el nombre, el contenido, la fecha de modificación y otra información sobre cada página wiki en el sistema. La opción de instalación predeterminada antepone `mdl_` delante de cada nombre de tabla (Wild, 2017). Cada componente principal del sistema suele tener una o más tablas, cada una de las cuales comienza con el nombre del componente (Moodle, 2020). Es importante poder tratar toda la base de datos como una entidad, copiando y moviendo instancias de una base de datos Moodle completa para la creación de áreas de prueba a medida que se desarrolla una herramienta o plugin.

Figura 7

Categorías que Guardan Relación en la Base de Datos de Moodle 3.9



Nota. Parte del esquema de la base de datos de Moodle. Adaptada de (Moodle, 2020).

Dada la extensa cantidad de tablas en la base de datos de Moodle, únicamente se mencionan las tablas que han sido relevantes para el estudio (Tabla 5). Mismas que permitirán la creación del dashboard de visualización.

Tabla 5

Descripción de las Tablas de Datos de Moodle Utilizadas para la Generación del Dashboard de Visualización

Tabla de Datos	Descripción
mdl_assign	Almacena información de los módulos de curso tipo assign (tarea) de los cursos.
mdl_assign_grades	Almacena las calificaciones de los estudiantes en relación a la actividad de tipo assign (tarea).
mdl_assign_submission	Almacena la fecha y el estado (BORRADOR/ENVIADO) de los envíos de una actividad tipo assign (tarea) que realizan los estudiantes.
mdl_assignsubmission_file	Almacena la cantidad de archivos que un estudiante envió en una actividad tipo assign (tarea).
mdl_course	Almacena información de cada una de las actividades y recursos que se encuentran en un curso. Contiene información sobre en qué curso y sección se muestra la actividad/recurso, así como detalles sobre la visibilidad, el grupo y el estado de finalización de la actividad.
mdl_course_sections	Almacena la información de las secciones que tiene cada curso.



mdl_forum	Almacena la información de los foros (forum) existentes en un curso.
mdl_grade_grades	Mantiene las calificaciones individuales para cada usuario y cada ítem de la tabla mdl_grade_items , exactamente cómo se importaron o enviaron los módulos.
mdl_grade_items	Contiene información sobre elementos calificables (es decir, columnas). Si una actividad (por ejemplo, una tarea o cuestionario) tiene varios elementos de calificación asociados (por ejemplo, varios resultados o calificaciones numéricas), entonces habrá un número múltiple correspondiente de filas en esta tabla.
mdl_logstore_standard_log	Almacena los registros de cada una de las interacciones que se producen en el sitio Moodle. Esta tabla se considera de vital importancia para el análisis de esta investigación. Puesto que mantiene registros de interacción del usuario con los módulos (actividades) o recursos del curso.
mdl_modules	Almacena información sobre los identificadores de los módulos (Sección 3.2.1) de Moodle disponibles para usarse dentro de un curso (assign , book , chat , choice , data , feedback , folder , forum , etc.).
mdl_question	Almacena las preguntas creadas para los módulos tipo quiz (cuestionario).
mdl_quiz	Almacena información relevante respecto a la actividad de tipo quiz (cuestionario) de un curso.
mdl_quiz_attempts	Almacena la información de los intentos de un usuario a un quiz (cuestionario).
mdl_quiz_slots	Almacena información de las preguntas y su orden en un quiz (cuestionario).
mdl_resource	Almacena información de los módulos de tipo resource (recurso) de un curso.
mdl_role	Guarda la información de los roles que pueden tener los usuarios en Moodle. Un rol es una colección de permisos definidos para todo el sistema que puede asignar a usuarios específicos en contextos específicos. La combinación de roles y contexto define la capacidad de un usuario específico para realizar alguna acción en el sitio. Los ejemplos más comunes son los roles de estudiante y docente en el contexto de un curso.
mdl_role_assignments	Guarda información sobre los roles que tiene un usuario dentro de un contexto de Moodle.
mdl_user	Esta tabla almacena información de los usuarios registrados en la plataforma, independientemente del tipo de rol que estén desempeñando.



Capítulo 4: Analíticas de Aprendizaje sobre Moodle

4.1. ¿Qué son las Analíticas de Aprendizaje?

Para estudiar la interacción de los estudiantes con estas plataformas han surgido áreas de investigación como la minería de procesos PM, la minería de procesos educativos EPM, la minería de datos DM, la minería de datos educativos EDM y analíticas de aprendizaje LA (Bogarín, Cerezo, & Romero, 2017). En la minería de procesos, la idea básica es extraer conocimiento de los registros de eventos registrados por un sistema de información (Bogarín, Romero, Cerezo, & Sánchez Santillán, 2014). La minería de procesos educativos se ha utilizado para extraer conocimiento de un tipo particular de un sistema de información educativa, considerando los procesos educativos que reflejan el comportamiento de los estudiantes (Luna, Castro, & Romero, 2017). La minería de datos es la extracción automática de patrones implícitos e interesantes de grandes colecciones de datos (Romero, Ventura, & Garcia, 2008). La minería de datos educativos se define como el desarrollo de métodos para explorar los tipos únicos de datos que provienen de entornos educativos, y usar esos métodos para comprender a los estudiantes y qué y cómo aprenden (Pardo, y otros, 2016). Por último, la analítica del aprendizaje que tiene como objetivo medir, recopilar y analizar datos educativos de las actividades de estudiantes y docentes para identificar patrones de comportamiento, comprender mejor proceso de enseñanza-aprendizaje y ayudar a mejorar los entornos en los que se producen estos procesos (Einhardt, Aires Tavares, & Cechinel, 2016). Aunque algunos autores hacen distinciones entre estas áreas debido al hecho de que su enfoque y rutas provienen de fuentes / lugares potencialmente diferentes (Luna, Castro, & Romero, 2017), estas áreas tienen el mismo objetivo de mejorar la calidad del análisis de datos educativos y enfrentan los mismos problemas.

4.2. Plugins de Analítica sobre Moodle

Hasta ahora, las comunidades EDM y LA han desarrollado una gran cantidad de herramientas para resolver problemas educativos (Romero & Ventura, 2016). Además, están surgiendo nuevas herramientas en diferentes entornos educativos LMS. Sin embargo, solo algunas de esas herramientas están específicamente orientadas para ser utilizadas con datos de Moodle (Luna, Castro, & Romero, 2017). Una de sus principales características de Moodle indicadas en su sitio web oficial, refiere a ser “altamente flexible y completamente personalizable” (Moodle, 2020). Haciendo énfasis en este aspecto, desencadena en la mejora continua de la experiencia de aprendizaje de los estudiantes mediante la personalización de cómo estos “miran y sienten el curso” (look and feel de Moodle). Las



herramientas específicas de Moodle se dividen en dos subtipos: herramientas de escritorio y basadas en la web. Las herramientas de escritorio son aplicaciones que se ejecutan localmente en un dispositivo informático. Las herramientas EDM y LA más antiguas que se usaron con los datos de Moodle fueron las herramientas de escritorio:

- **CoSyLMSAnalytics:** Es un ejemplo de estas herramientas, que permite que un docente evalúe y supervise el progreso individual o colectivo, a fin de proporcionar información a los estudiantes mediante informes. Recopila información sobre los patrones de acceso de los alumnos y extrae correlaciones y reglas de asociación entre sus rutas de aprendizaje (Retalis, Papasalouros, Psaromiligkos, Siscos, & Kargidis, 2006).
- **ViMoodle:** Proporciona el beneficio adicional de comprender las relaciones educativas en Moodle. Su objetivo principal es extraer conocimiento específico relacionado con el complejo proceso de educación y aprendizaje. Utiliza actividades de registro, cuestionarios, foros, recursos, etc., de la base de datos y devuelve gráficos de redes sociales y caracoles utilizando algoritmos de minería de reglas de asociación y agrupamiento (Gómez Aguilar, Therón, & García Peñalvo, 2008).
- **CIECoF o Marco de curso de mejora continua del aprendizaje electrónico:** proporciona recomendaciones a los autores de cursos sobre cómo mejorar los cursos de Moodle. Su objetivo es encontrar, compartir y sugerir las modificaciones más apropiadas para mejorar la efectividad del curso. Utiliza la minería de reglas de asociación y el filtrado colaborativo a través de los datos de uso de los estudiantes (García, Romero, Ventura, & Castro, 2008).
- **Meerkat-ED:** herramienta que permite analizar la participación de los estudiantes en los foros de discusión de Moodle. Su objetivo es proporcionar al instructor mejores medios para evaluar la participación en debates en línea. Utiliza técnicas de análisis de redes sociales sobre los datos de discusión del foro (Rabbany, Takaffoli, & Zaïane, 2011).
- **MMT o Moodle-Mining Tool:** es una herramienta diferente que hace que la ejecución de todos los pasos en el proceso de minería de datos. Su objetivo es proporcionar técnicas de DM no solo para expertos sino también para los recién llegados. Aplica algoritmos de minería de reglas de preprocesamiento, clasificación y asociación sobre archivos de datos creados a partir de la base de datos Moodle (Pedraza Pérez, Romero, & Ventura, 2011).
- **DRAL:** revela actividades relevantes para los alumnos de Moodle. Su objetivo es detectar las actividades que un estudiante necesita para aprobar un curso en función de las características



extraídas de los datos registrados. Utiliza programación genética de objetivos múltiples y gramatical para tareas de clasificación (Zafra, Romero, & Ventura, 2013).

Las herramientas basadas en la web se entregan a un dispositivo local a través de Internet desde un servidor remoto. Actualmente, casi todas las nuevas herramientas EDM y LA están basadas en la web y a continuación se describen algunos ejemplos.

- **GISMO:** utiliza el monitoreo gráfico interactivo de los estudiantes para extraer datos de seguimiento de un curso en línea mantenido con Moodle. Su objetivo es proporcionar visualizaciones de datos conductuales, cognitivos y sociales del curso, permitiendo un monitoreo constante de las actividades, el compromiso y los resultados de aprendizaje de los estudiantes. GISMO está integrado como un bloque en Moodle (Mazzola, Nidola, Milani, & Mazza, 2010).
- **SNAPP o Redes Sociales Adaptando la Práctica Pedagógica:** es una herramienta que visualiza la evolución de las relaciones de los participantes en los foros de discusión. Su objetivo es analizar los patrones de interacción emergentes para permitir que se realicen las intervenciones según sea necesario. Está integrado en la interfaz de Moodle como Bookmarklet, pero solo muestra la visualización en tiempo real de la actividad del foro de discusión (Bakharia & Dawson, 2011).
- **AAT o Academic Analytics Tool:** accede y analiza los datos de comportamiento de los estudiantes en los sistemas de aprendizaje. Su objetivo es permitir la identificación de material de aprendizaje difícil o inapropiado con el fin de contribuir significativamente al diseño de actividades y recursos mejorados de apoyo al estudiante. Aunque está integrado en Moodle como un bloque, genera representaciones gráficas que solo pueden ser exploradas por los docentes del curso (Graf, Ives, Rahman, & Ferri, 2011).
- **MOClog:** es una herramienta para el análisis y la presentación de datos de registro en un servidor Moodle. Su objetivo es medir el estado de las actividades en el curso en línea tanto como sea posible confiando exclusivamente en los datos del archivo de registro. Reutiliza los componentes disponibles de GISMO (Mazzola, Nidola, Milani, & Mazza, 2010) para contar y visualizar el número de botes y clics (Mazza, Bettoni, Faré, & Mazzola, 2012).
- **E-learningWebMiner:** revela los perfiles de comportamiento de los estudiantes y modela cómo trabajan en cursos virtuales. Su objetivo es proporcionar información útil que los



docentes pueden usar para mejorar sus cursos. Es un servicio web que proporciona gráficos de visualización, agrupación y algoritmos de asociación (Zorrilla & García Saiz, 2013).

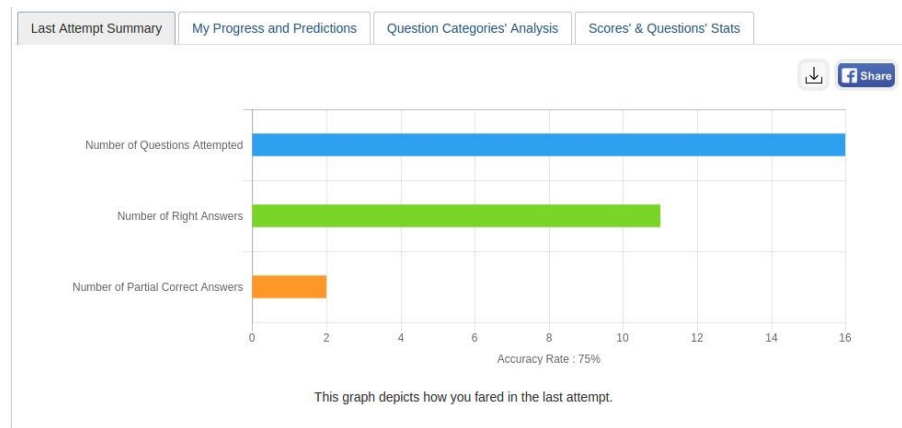
- **CVLA:** integra técnicas analíticas en un informe personalizado de Moodle. Su objetivo es utilizar múltiples conjuntos de datos y técnicas analíticas en una sola interfaz para presentar datos a estudiantes y docentes. Está integrado en Moodle como un módulo y proporciona análisis de redes sociales y algoritmos de clasificación para predecir el envío de tareas (Drăgulescu, Bucos, & Vasiiu, 2015).
- **SmartKlass:** mide y analiza el proceso de aprendizaje en cualquier momento a lo largo de los cursos de Moodle. Su objetivo es capacitar a los docentes para gestionar la evolución de los estudiantes en un curso en línea. Es un plugin del panel de análisis de aprendizaje de código abierto y multiplataforma (KlassData Team, 2016).
- **MEAP o Moodle Engagement Analytics Plug-in:** proporciona comentarios sobre el nivel de participación de un estudiante en un curso de Moodle. Su objetivo es identificar actividades que tengan un impacto en el éxito del estudiante. Está integrado como un bloque en Moodle (Yen-Ting Liu, Froissard, & Richards, 2015).

La búsqueda de extensiones (plugins) para el dashboard propuesto que mejoran el look and feel o permitan una mejor visualización de analíticas de aprendizaje en el curso, requiere de extensiones compatibles con la versión 3.X de Moodle, considerando que hasta la fecha es la versión más actual del LMS. A continuación, se incluye una breve descripción de esas herramientas actuales.

- **Grade Reports – Quiz Analytics:** Este complemento proporciona a los usuarios analíticas respecto al módulo de tipo quiz (cuestionarios/evaluaciones). Cada intento que cada usuario hace en un cuestionario se analiza y los datos relevantes se representan a través de gráficos visualmente atractivos (DualCube Team, 2020). Los gráficos se distribuyen en las siguientes pestañas:
 - **Resumen de intentos:** esta pestaña contiene las estadísticas del usuario que inició sesión para el último intento que realizó en ese cuestionario. En caso de pruebas que permitan un solo intento, ese intento se considera el último intento (Figura 8).

Figura 8

Descripción del Último Intento en una Evaluación/Cuestionario



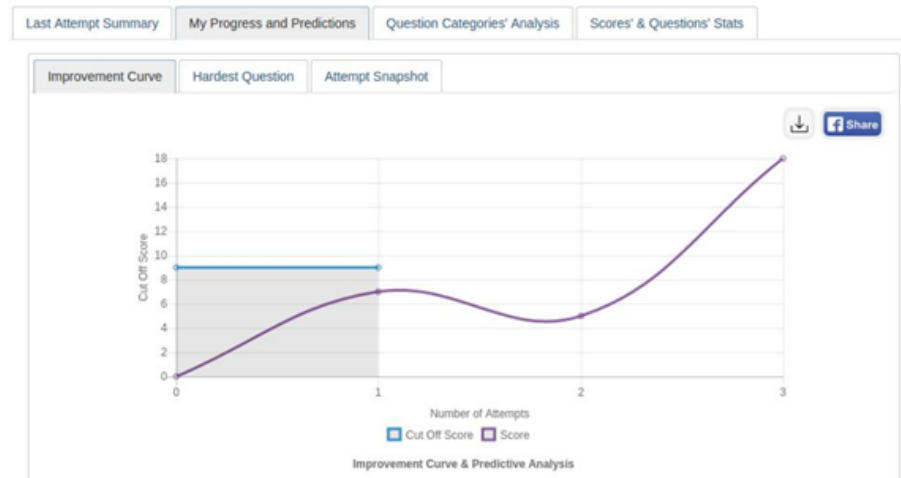
Nota. Tomada de (DualCube Team, 2020).

- **Mi progreso y predicciones:** esta pestaña contiene varias subpestañas que, en su conjunto, dan una idea de cómo le ha ido al usuario que inició sesión con el tiempo y la pregunta más difícil que enfrentó en la evaluación/cuestionario. También hay una pestaña específica para cuestionarios que ofrece múltiples intentos para mostrar al usuario registrado un análisis predictivo en función de los intentos y mejores puntuaciones de otros usuarios (Figura 9, Figura 10, Figura 11, Figura 12).



Figura 9

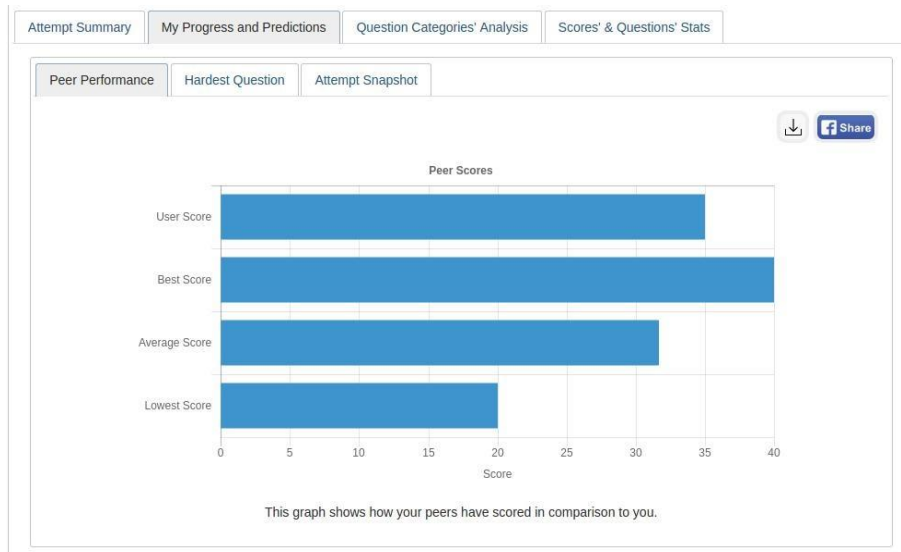
Curva de Mejora en la Pestaña de Mi Progreso y Predicciones



Nota. Muestra cómo mejoró en todos sus intentos y el bloque de la zona oscura configurado por el usuario es el umbral de aprobación. Además, el eje y representa los intentos. Tomada de (DualCube Team, 2020).

Figura 10

Desempeño de los Compañeros en la Pestaña de Mi Progreso y Predicciones



Nota. Representa de forma global una comparación del curso en esa evaluación contra el rendimiento individual del usuario. Tomada de (DualCube Team, 2020).



Figura 11

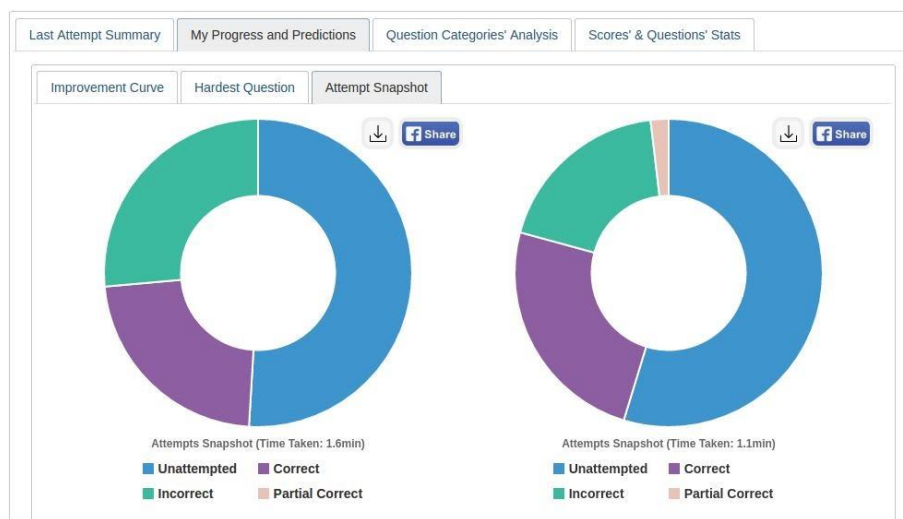
Pregunta más Difícil en la Pestaña de Mi Progreso y Predicciones



Nota. Muestra cuántas veces se intentó responder una pregunta de forma incorrecta, si se da click en la barra púrpura hay una explicación de la respuesta correcta. Tomada de (DualCube Team, 2020).

Figura 12

Resumen de Intentos en la Pestaña de Mi Progreso y Predicciones



Nota. Resumen que muestra las cifras clave de todos sus intentos anteriores. Tomada de (DualCube Team, 2020).

- **Análisis de categorías de preguntas:** esta pestaña está dedicada a las categorías/etiquetas con las que están asociadas todas las preguntas. La Figura 13 muestra la distribución de preguntas entre categorías y también una idea de cuán desafiante resultó ser cada una de las categorías.

Figura 13

Categorías Desafiantes para Mí en la Pestaña de Análisis de Categoría de Preguntas



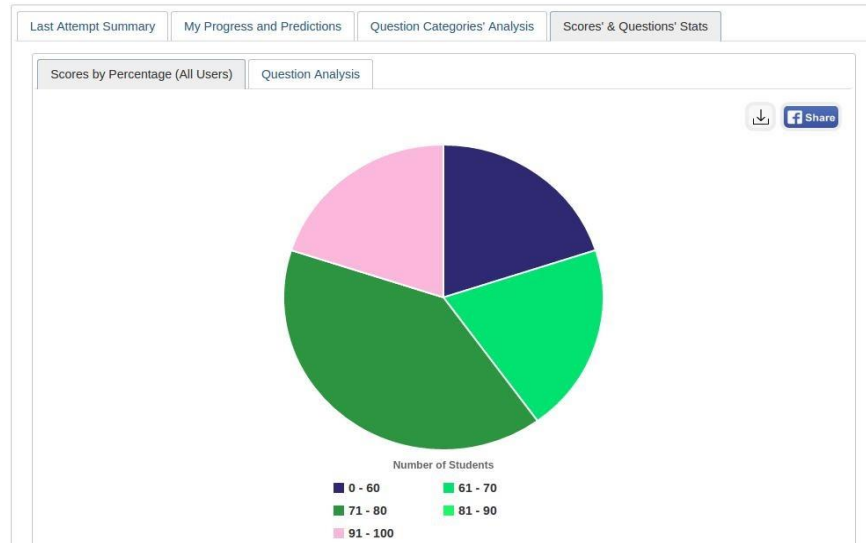
Nota. Gráfica de las categorías de bajo rendimiento en la evaluación/cuestionario. Tomada de (DualCube Team, 2020).

- **Estadísticas de puntuaciones y preguntas:** esta pestaña contiene dos sub pestañas, una que muestra el número de estudiantes en cada categoría de grado (Figura 14) y la otra que proporciona un análisis exhaustivo de todas las preguntas presentes en el cuestionario (Figura 15).



Figura 14

Calificaciones por Porcentajes en la Pestaña de Estado de Pregunta y Calificaciones



Nota. Representa el número de usuarios por cada porcentaje. Tomada de (DualCube Team, 2020).

Figura 15

Análisis de Preguntas en la Pestaña de Estado de Preguntas y Calificaciones



Nota. Las curvas representan cómo fue el desempeño de los usuarios en cada pregunta, en cada círculo se puede dar click y ver más detalles de la pregunta. Tomada de (DualCube Team, 2020).

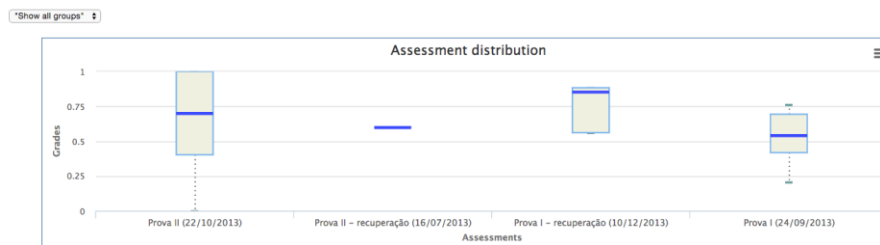


- **Blocks: Analytics graphs:** El plugin es una herramienta de análisis de aprendizaje de nivel inferior. Recopila datos de la actividad de los estudiantes existentes de Moodle y los muestra en varios gráficos. En otras palabras, estos datos siempre han estado disponibles para los profesores en Moodle, pero no en forma visual fácilmente visualizable. Analytics Graphs ayuda a entender los datos recopilados de los registros del curso y las actividades calificadas. Esta herramienta ofrece a los docentes una variedad de gráficos con los que examinar la actividad de los estudiantes en su curso. Cada cuadro se puede ver para un grupo de estudiantes específico (si los estudiantes están inscritos en grupos de Moodle) y se puede descargar o imprimir (Schmitt, 2018). Esta herramienta genera gráficos destinados a facilitar decisiones pedagógicas. Los gráficos tienen capacidades de zoom y permiten una comunicación rápida con los estudiantes a través del correo electrónico. Si se usa de manera eficiente, el complemento Analytics Graphs puede ser útil para aumentar la participación de los estudiantes en el estudio y, en consecuencia, para reducir las tasas de deserción (Schmitt, 2018). Este plugin proporciona cinco gráficos que facilitan la identificación de los perfiles de los estudiantes, lo que permite al maestro enviar mensajes a los usuarios de acuerdo con su comportamiento dentro de un curso (Schmitt, 2018). Los gráficos disponibles se listan a continuación.
 - **Cuadro de calificaciones:** muestra la distribución de calificaciones en un gráfico de caja para identificar las diferencias entre las evaluaciones y los estudiantes con problemas (Figura 16). Esta página obtiene resultados de calificaciones de tareas, cuestionarios, bases de datos, foros, cuestionarios de puntos, etc., pero solo si los envíos de los estudiantes en estas actividades han sido calificados (ya sea automáticamente por Moodle o manualmente por un docente).



Figura 16

Tabla de Calificaciones, Representación con Diagrama de Cajas y Bigotes

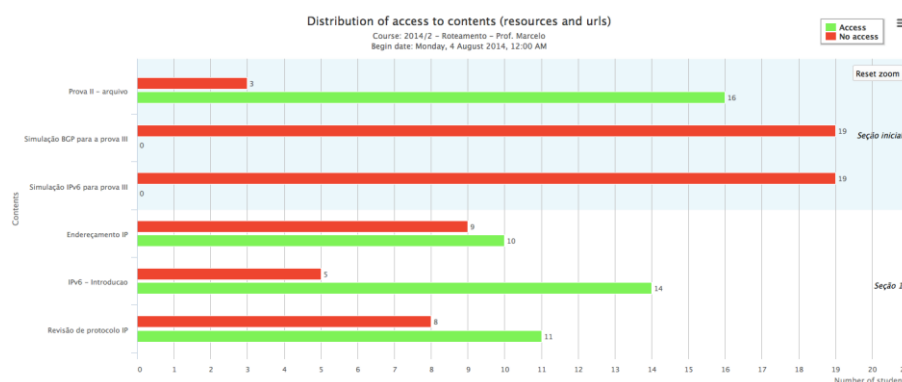


Nota. Tomada de (Schmitt, 2018).

- **Tabla de accesos al contenido:** qué usuarios accedieron a los diferentes recursos que el docente pone a disposición de los estudiantes (Figura 17). Esta página puede enumerar todas las actividades y recursos en la página del curso (incluidos los que están ocultos a la vista de los estudiantes) e indicar cuántos estudiantes han accedido a alguno de ellos.

Figura 17

Representación Gráfica del Acceso al Contenido del Curso



Nota. Tomada de (Schmitt, 2018).

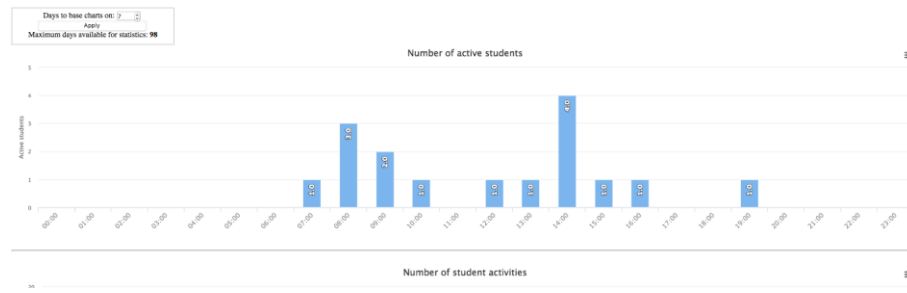
- **Gráfico de número de usuarios activos:** cómo los usuarios muestran actividad en la plataforma en un momento determinado del día (Figura 18). Es posible hacer clic



sobre los elementos del gráfico para enviar un correo electrónico a un grupo de estudiantes o a un estudiante en particular.

Figura 18

Gráfico de Usuarios Activos

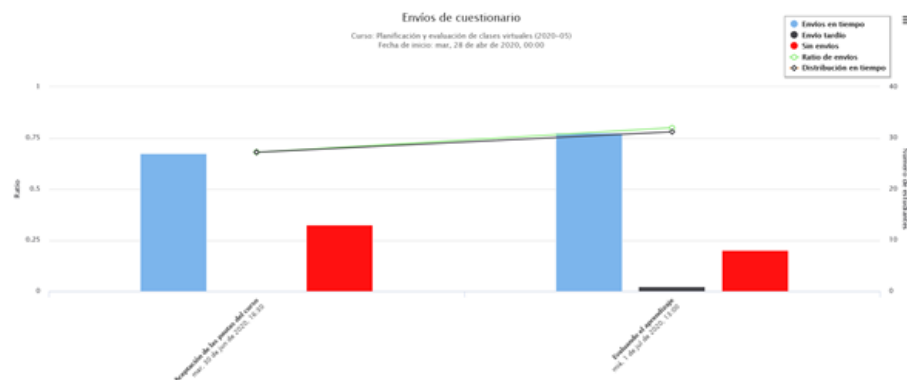


Nota. Tomada de (Schmitt, 2018).

- **Tabla de envíos de tareas:** qué usuarios han enviado las tareas a tiempo o tarde (tareas, cuestionarios o alguna actividad evaluable) (Figura 19). Esta página puede enumerar todas las tareas de un curso en el orden en que aparecen en la página del curso. Debajo de la tabla se pueden seleccionar las asignaciones del curso para su visualización. Un gráfico de barras indica por tarea cuántos estudiantes entregaron la tarea antes de la fecha límite, cuántos se entregaron tarde y cuántos no.

Figura 19

Cuadro de Envío de Tareas o Cuestionarios



Nota. Tomada de (Schmitt, 2018).



- **Cuadro de distribución de visitas:** cómo cada usuario accede al curso y sus recursos en cada semana del curso (Figura 20).

Figura 20

Distribución de Visitas al Curso

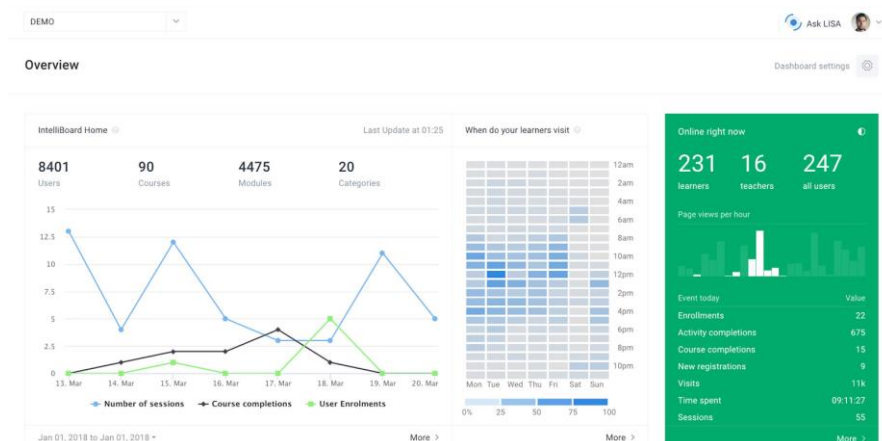


Nota. Tomada de (Schmitt, 2018).

- **IntelliBoard.net:** extrae los datos estadísticos recopilados y disponibles en Moodle y los presenta en un tablero/dashboard en forma de cuadros, gráficos e informes de formato múltiple imprimibles (Figura 21). Su objetivo es entregar análisis de datos educativos a un solo tablero al instante. Es un servicio web analítico y de informes comerciales y un plugin (IntelliBoard Team, 2020).

Figura 21

Dashboard de IntelliBoard



Nota. Se ha hecho uso de una demo de esta herramienta. Tomada de (IntelliBoard Team, 2020).



- **Learnerscript:** es un plugin específico de Moodle que se puede instalar localmente. Es decir, la instalación se realiza en el servidor que contenga Moodle. Contiene cuatro dashboards para el administrador. Los roles de maestro y alumno tienen un dashboard cada uno. Los gráficos y tablas disponibles miden y analizan los resultados de aprendizaje. Ofrece datos con respecto a recuento diario de visitantes al sitio, número de visitas, tiempo promedio empleado, tiempo total dedicado, registro de dispositivos de acceso al sitio, informes relacionados cursos, tareas, cuestionarios, etc. (LearnerScript Team, 2020).

4.3. Patrones de Aprendizaje en Plataformas Virtuales

El campo emergente de análisis de aprendizaje está desarrollando constantemente nuevas formas de analizar datos sobre las interacciones, el compromiso y rendimiento de los estudiantes (Coffrin, Corrin, de Barba, & Kennedy, 2014). Las analíticas del aprendizaje se han centrado en apoyar a los docentes en realizar un seguimiento del progreso de los estudiantes a lo largo del tiempo a partir de fuentes de datos, interacciones y comportamientos de estudio (Ginda, Suri, Bueckle, & Börner, 2016). Al examinar las huellas digitales de los alumnos, algunas investigaciones han encontrado útil identificar patrones de participación de los alumnos, que generalmente están relacionados con su compromiso y logro (Coffrin, Corrin, de Barba, & Kennedy, 2014). La minería de datos se puede aplicar para explorar, visualizar y analizar datos de aprendizaje (Mazzola, Nidola, Milani, & Mazza, 2010) con el fin de identificar patrones útiles (Gaudioso & Talavera Méndez, 2005), evaluar la actividad web y para obtener más información sobre cómo aprenden los alumnos.

Aunque el uso de LMS se ha generalizado en el contexto de la educación superior, áreas como LA y EDM se enfocan en analizar datos de estos sistemas, sin embargo, se ha prestado poca atención a la exploración de comportamientos únicos de los estudiantes en entornos de aprendizaje virtual (Juhaňák, Zounek, & Rohlíková, 2019). La minería de procesos tiene sus raíces en los negocios, pero como se mencionó anteriormente, está comenzando a usarse en el contexto educativo. La minería de procesos se refiere a un conjunto de métodos, técnicas y herramientas que se centran en extraer conocimiento relacionado con el proceso de los llamados registros de eventos, que comúnmente se almacenan en diferentes tipos de sistemas de información, incluidos los educativos del tipo LMS (Sonnenberg & Bannert, 2015).



4.4. Predicción en Base a Interacciones de Aprendizaje

En los últimos años la cantidad de información ha sido notable en los campos de análisis de aprendizaje y minería de datos educativos dedicada al modelado predictivo (Pardo, y otros, 2016). Los modelos comúnmente ayudan a identificar posibles relaciones entre varios factores conductuales, demográficos y basados en el rendimiento en grandes conjuntos de datos (Ginda, Suri, Bueckle, & Börner, 2016). Un ejemplo radica en la investigación asociada con la identificación temprana de estudiantes en riesgo de rendimiento académico o desgaste (Pardo, y otros, 2016). Otro aspecto a tener en cuenta son las altas tasas de deserción de los alumnos. Si bien estos tipos de modelos predictivos se han utilizado bien para desarrollar indicadores tempranos, se ha prestado poca atención a la investigación de cómo dicha información se puede implementar mejor para promover la reflexión y la acción entre el personal docente y los estudiantes (G. de Barba, y otros, 2020).

4.5. Analíticas de Aprendizaje y Dashboards

Dentro de los diversos aspectos que conforman el área de analíticas de aprendizaje, la visualización de datos educativos se ha considerado un área crítica que merece atención urgente por parte de los investigadores (Einhardt, Aires Tavares, & Cechinel, 2016). El gran desafío actual es cómo los datos provenientes de las plataformas de aprendizaje pueden ser procesables analizándolos y presentándose de manera significativa para los diferentes interesados (Abubakar, Kamsin, & Abdullah, 2020). Ya que como señala Schwendimann, aunque es visualmente atractivo, muchos paneles no tienen la capacidad de proporcionar información realmente útil (Schwendimann, Rodríguez Triana, Vozniuk, Prieto, & Shirvani Boroujeni, 2017). En este contexto, las herramientas de visualización de datos educativos son consideradas como una de las brechas que existen actualmente en el área de analíticas del aprendizaje (Einhardt, Aires Tavares, & Cechinel, 2016).

Los dashboards o paneles se utilizan constantemente en todas las áreas y sistemas para presentar visualmente a través de gráficos un resumen de lo que se observa en el sistema (D. Roberts, A. Howell, & Seaman, 2017). Smith definió un panel de información como "una presentación visual de la información más importante necesaria para lograr uno o más objetivos; consolidado y organizado en una sola pantalla para que la información pueda ser monitoreada de un vistazo" (Smith, 2013). Schwendimann describió los paneles como "una interfaz de usuario en tiempo real, fácil de leer, a menudo de una sola página, que muestra una presentación gráfica del estado actual (instantánea) y las tendencias históricas de los Indicadores Clave de Rendimiento (*KPI – Key Performance Indicator en inglés*) de una organización para permitir que se tomen decisiones



instantáneas e informadas hecho de un vistazo" (Schwendimann, Rodríguez Triana, Vozniuk, Prieto, & Shirvani Boroujeni, 2017). Para que los paneles de instrumentos sean útiles como herramientas de apoyo para la toma de decisiones y el aprendizaje, la investigación tiene el desafío de identificar qué y cómo se muestra la información en los paneles de manera oportuna y precisa (Einhardt, Aires Tavares, & Cechinel, 2016).



Capítulo 5: Conceptos Relacionado a Datos

En la actualidad la cantidad de datos educativos generados por diversos LMS, han representado una especie de “caja negra” (Juhaňák, Zounek, & Rohlíková, 2019). Aunque en muchos casos, estos sistemas recopilan y almacenan información sobre su uso en forma de registros generados automáticamente en un servidor o en la base de datos del sistema, estos datos rara vez se emplean en la investigación o la institución los utiliza sistemáticamente (Shaun Bake & Salvador Inventado, 2014). Debido al tamaño y complejidad de los datos que estos sistemas LMS manejan, extraer información útil y relevante se vuelve una tarea compleja por lo que son necesarios métodos de extracción, análisis, limpieza y carga de información en una única fuente de datos (Codish, Rabin, & Ravid, 2019), que facilite la obtención de información con el nivel semántico necesario para responder una pregunta de investigación en particular (Valle Torre, Tan, & Hauff, 2020).

5.1. Datos, Información y Conocimiento

5.1.1. Datos

Se consideran elementos primarios de la información cuyo valor por sí solos no son de relevancia para la toma de decisiones, son la mínima unidad semántica; los datos pueden ser una colección de hechos almacenados en algún dispositivo; pueden ser internos o externos a la organización, objetivos o subjetivos o de tipo cualitativos y cuantitativos (Chen, y otros, 2009). Los datos son símbolos que describen condiciones, hechos, situaciones o valores. Los datos se caracterizan por no contener ninguna información. Un dato puede significar un número, una letra, un signo ortográfico o cualquier símbolo que represente una cantidad, una medida, una palabra o una descripción. (Zins, 2007).

5.1.2. Información

La información está constituida por un grupo de datos ya supervisados y ordenados, que sirven para construir un mensaje basado en un cierto fenómeno o ente. La información permite resolver problemas y tomar decisiones, ya que su aprovechamiento racional es la base del conocimiento (Kreulen, Cody, Krishna, & Spangler, 2002).

(Zins, 2007) afirma que la información consiste en un conjunto de datos que poseen un significado, de modo tal que reducen la incertidumbre y aumentan el conocimiento de quien se acerca a contemplarlos. Estos datos se encuentran disponibles para su uso inmediato y sirven para clarificar incertidumbres sobre determinados temas. Por otro lado, (Shaun Bake & Salvador Inventado, 2014)



indican que la información es un conjunto de datos procesados con un significado, un propósito y/o contexto; cuya utilidad reside en la toma de decisiones debido a la minimización de la incertidumbre.

5.1.3. Conocimiento

El conocimiento es un conjunto de información almacenada mediante la experiencia o el aprendizaje (a posteriori), o a través de la introspección (a priori). En el sentido más amplio del término, se trata de la posesión de múltiples datos interrelacionados que, al ser tomados por sí solos, poseen un menor valor cualitativo (Zins, 2007). (Chen, y otros, 2009) establecen la definición de conocimiento como datos que representan los resultados de un proceso cognitivo simulado por computadora, como la percepción, el aprendizaje, la asociación y el razonamiento, o las transcripciones de algunos conocimientos adquiridos por los seres humanos.

5.2. Inteligencia de Negocios (*BI – Business Intelligence en inglés*)

Business Intelligence (BI) es la habilidad para transformar los datos en información, y la información en conocimiento, de forma que se pueda optimizar el proceso de toma de decisiones en los negocios (Foley & Guillemette, 2010). Desde un punto de vista asociado con las ICT, se puede definir Business Intelligence como el conjunto de metodologías, aplicaciones y tecnologías que permiten reunir, depurar y transformar datos de los sistemas transaccionales e información desestructurada (interna y externa a la compañía) en información estructurada, para su análisis y conversión en conocimiento (M. Müller & Joachim Lenz, 2013), dando así soporte a la toma de decisiones sobre el negocio.

5.3. Data Warehouse

Un Data Warehouse es una base de datos corporativa que se caracteriza por integrar y depurar información de una o más fuentes distintas, para luego procesarla permitiendo su análisis desde infinidad de perspectivas y con grandes velocidades de respuesta (J. Watson & H. Wixom, 2007). La creación de un datawarehouse representa en la mayoría de las ocasiones el primer paso, desde el punto de vista técnico, para implantar una solución completa y fiable de Business Intelligence (M. Müller & Joachim Lenz, 2013).

Es una colección de datos orientada a un determinado ámbito (empresa, organización, etc.), integrado, no volátil y variable en el tiempo, que ayuda a la toma de decisiones en la entidad en la que se utiliza (Bustos Barrera & Mosquera Artieda, 2013). Una de las definiciones más famosas sobre DWH, es la de William Harvey Inmon, quien define: “Un Data Warehouse es una colección de datos



orientada al negocio, integrada, variante en el tiempo y no volátil para el soporte del proceso de toma de decisiones de la gerencia” (Foley & Guillemette, 2010). Se trata de un expediente completo de una empresa, más allá de la información transaccional y operacional, almacenada en una base de datos, diseñada para favorecer el análisis y la divulgación eficiente de datos (Peña Rivera & Suárez Daza, 2008).

5.4. Data Warehousing

El Data Warehousing, es el término que se utiliza para describir la creación, el mantenimiento, el uso y la actualización continua de los datos (Watson, Ariyachandra, & Matyska, 2001). (Abril Lara, 2016) considera que el Data Warehousing es el proceso encargado de extraer, transformar, consolidar, integrar y centralizar los datos que una organización genera con su actividad diaria. Este proceso permite el acceso y exploración de la información requerida, a través de una amplia gama de posibilidades de análisis, con el objetivo final de dar soporte al proceso de toma de decisiones estratégico y táctico (Golfarelli, Rizzi, & Cella, 2004).

Por otro lado, (Silva Peñafiel, 2018) sugiere que la implantación de almacenes de datos (Data Warehouse) ha generado diferentes técnicas de gran alcance, que permiten extraer datos operativos desde diferentes formatos, así como la integración de los datos, independientemente de su ubicación, formato o requisitos de comunicación; por lo que se puede convertir un almacén de datos “inactivo” en una fuente de información “dinámica”; a esto lo denominamos Data Warehousing.

5.5. ETL (Extract, Transform and Load)

Un Data Warehouse o Data Mart, se carga periódicamente, y en él se unifica información procedente de múltiples fuentes. Esto implica que deben existir una serie de procesos que leen los datos de las diferentes fuentes, los transforman y adaptan al modelo definido, los depuran y limpian, y los introducen en esta base de datos de destino (Calvanese, Giacomo, Lenzereni, Nardi, & Rosati, 2001). Esto es lo que se conoce como procesos ETL, procesos de Extracción, Transformación y Carga (Load) (Figura 22).

Para (Gonzales Farro, 2018), el proceso ETL es de vital importancia, ya que antes de almacenar los datos en un Data Warehouse, éstos deben ser transformados, limpiados, filtrados y redefinidos, ya que la información que se tiene en los sistemas transaccionales no está preparada para la toma de decisiones.

Figura 22*Proceso de Extracción, Transformación y Carga de datos*

Nota. Adaptada de (Calvanese, Giacomo, Lenzereni, Nardi, & Rosati, 2001).

5.6. Arquitectura del Data Warehousing

La arquitectura está formada por diferentes elementos que interactúan entre sí para cumplir una función en específico dentro de un sistema. Pero se debe siempre realizar un proceso para que las operaciones sean comprensibles y eficientes (Abril Lara, 2016), como se muestra en la Figura 23.

Figura 23*Arquitectura del Data Warehousing*

Nota. Adaptada de (Bernabeu, 2009).

Los datos son extraídos desde aplicaciones, bases de datos, archivos, etc. Esta información generalmente reside en diferentes tipos de sistemas, orígenes y arquitecturas y tienen formatos muy variados (Ramos, 2011). Los datos son integrados, transformados y limpiados, para luego ser cargados en el Data Warehouse. Principalmente, la información del DW se estructura en bases de



datos multidimensionales, ya que estos preparan esta información para responder a consultas dinámicas con un buen rendimiento (Gonzales Farro, 2018). Los usuarios acceden a estas bases de datos multidimensionales o también denominados cubos multidimensionales (u otro tipo de estructura de datos) del DW utilizando diversas herramientas de consulta, exploración, análisis, reportes, etc. (Bernabeu, 2009).

5.7. Modelos y Bases de Datos Multidimensionales

El modelo multidimensional es una representación lógica de un proceso empresarial cuyas características importantes son la comprensión del usuario, el rendimiento de las consultas y la resistencia al cambio (Song, Rowen, Medsker, & Ewen, 2001). Una base de datos multidimensional es una base de datos en donde su información se almacena a través de tablas de hechos y tablas de dimensiones (Abril Lara, 2016). Proveen una estructura que permite tener acceso flexible a los datos, para explorar y analizar sus relaciones, y resultados (Bernabeu, 2009).

5.8. Metodología Hefesto

En esta sección se presenta la metodología Hefesto, que permitirá la construcción de un Data Warehouse de manera sistemática e intuitiva. Hefesto es una metodología propia, cuya propuesta está fundamentada en una muy amplia investigación, comparación de metodologías existentes y experiencias propias en procesos de confección de almacenes de datos (Bernabeu, 2009).

Hefesto es una metodología cuya finalidad es guiar el modelado del Data Warehouse de manera sencilla, ordenada y de forma intuitiva (Gonzales Farro, 2018). Con la ayuda de esta metodología es posible obtener, en poco tiempo, una visión integral del problema y del modelo constructivo, suficientemente sólido, para desarrollar una implementación completa del sistema requerido (Bustos Barrera & Mosquera Artieda, 2013). Para llevar a cabo la construcción de un almacén de datos, esta metodología sugiere distribuir el proceso en cuatro fases (Silva Peñafiel, 2018).

Hefesto propone, en la primera etapa, identificar las necesidades de los usuarios para conocer las preguntas claves del negocio. Luego, se continúa con el diseño de los indicadores que van a responder a estos interrogantes junto con las perspectivas de análisis requeridas (Bustos Barrera & Mosquera Artieda, 2013). Más tarde, se construye el modelo conceptual de datos del Data Warehouse. En etapas posteriores, se analizan las fuentes de datos para determinar los factores constructivos de cada indicador y se señalan las correspondencias con las fuentes de datos (Gonzales Farro, 2018). Una vez hecho esto, se podrá continuar con la construcción del modelo lógico del Data Warehouse. En esta etapa, deberá quedar definido el tipo de esquema utilizado en la implementación. También,



se construirán las tablas de dimensiones y las tablas de hechos para, finalmente, efectuar las uniones correspondientes. Por último, se definirán los procesos de extracción, transformación y carga para capturar los datos desde las diferentes fuentes de datos y, finalmente, almacenarlos en el Data Warehouse (Bernabeu, 2009).

5.8.1. Análisis de requerimientos

El análisis de requerimientos se centra en la identificación de las necesidades del cliente, esto puede desarrollarse mediante preguntas que ayuden a la identificación de los objetivos que persigue la organización (Silva Peñafiel, 2018); para luego determinar los indicadores y perspectivas en las cuales se basarán el modelo del Data Warehouse (Calvanese, Giacomo, Lenzereni, Nardi, & Rosati, 2001).

5.8.1.1. Identificar preguntas

Lo fundamental en este paso es formular preguntas sobre la organización, las cuales integran variables de análisis que sean relevantes para el estudio de la información desde diferentes perspectivas (Abril Lara, 2016). Hay que tener presente, que la información obtenida debe estar soportada de alguna manera por algún Sistema de Procesamiento de Transacciones en Línea (*OLTP – OnLine Transaction Processing en inglés*) (Ramos, 2011), ya que de otra manera no se podría implementar el Data Warehouse.

5.8.1.2. Identificar indicadores y perspectivas de análisis

La identificación de los indicadores y perspectivas se deriva de una descomposición de las preguntas obtenidas anteriormente, entre los indicadores efectivos se encuentran aquellos valores numéricos que representan lo que se desea analizar (Abril Lara, 2016), por ejemplo: en una empresa de producción y ventas; los saldos, promedio de ventas, cantidades vendidas; mientras que las perspectivas se refieren a los objetos mediante los cuales se necesitan analizar los indicadores (Silva Peñafiel, 2018), por ejemplo: clientes, proveedores, sucursales, productos, entre estos identificamos también el tiempo.

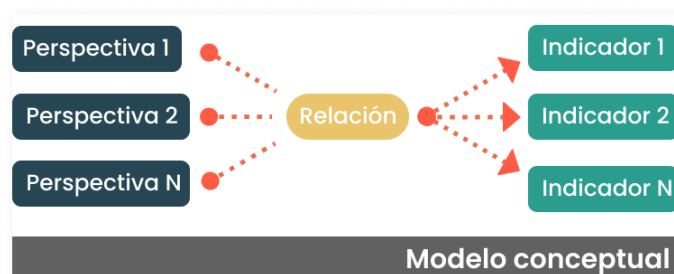
5.8.1.3. Modelo Conceptual

Un modelo conceptual consiste en la descripción de alto nivel de la estructura de la base de datos, en la cual la información se presenta a través de objetos, relaciones y atributos (Ramos, 2011). La construcción de este modelo se realiza en base a los indicadores y perspectivas (Gonzales Farro,

2018), las cuales nos permitirán observar de una forma clara y precisa los alcances del proyecto (Figura 24).

Figura 24

Estructura General de un Modelo Conceptual



Nota. Adaptada de (Bernabeu, 2009).

5.8.2. *Análisis de los OLTP*

El análisis de las fuentes de datos OLTP permite establecer cómo se obtendrán y calcularán los indicadores de acuerdo con el modelo conceptual creado anteriormente y las fuentes de datos que se disponen (Bustos Barrera & Mosquera Artieda, 2013); luego se incluirán en los campos para cada perspectiva y finalmente se ampliará el modelo conceptual antes definido (Silva Peñafiel, 2018).

5.8.2.1. **Determinar indicadores**

El cálculo de los indicadores debe ser detallado en este paso, de acuerdo con los conceptos siguientes (Silva Peñafiel, 2018):

- Los hechos que componen los indicadores y su fórmula para el cálculo (Bernabeu, 2009), como, por ejemplo:

$$\text{indicador} = \text{hecho1} + \text{hecho2}$$

- Función de summarización que se utilizará para la agregación, tales como *sum*, *avg*, *count* (Bustos Barrera & Mosquera Artieda, 2013).

5.8.2.2. **Establecer correspondencias**

En esta etapa, se requiere identificar las correspondencias entre el modelo conceptual formulado y las fuentes de datos obtenidas de los OLTP disponibles (Silva Peñafiel, 2018). El

objetivo es, que todos los elementos del modelo conceptual estén correspondidos en los OLTP (Abril Lara, 2016).

5.8.2.3. Nivel de granularidad

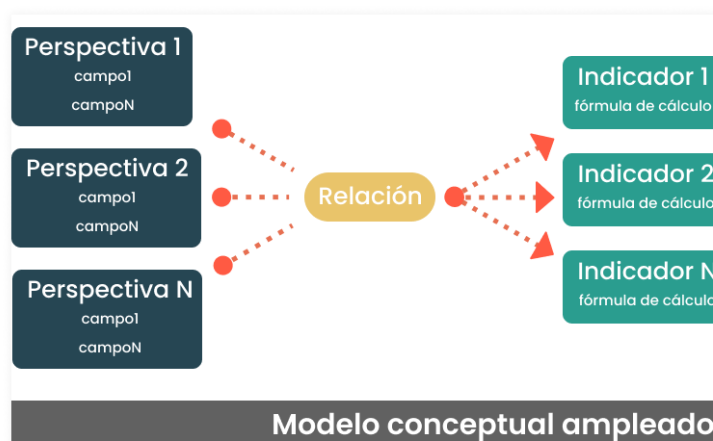
En esta etapa se procede a la selección de los campos que contendrá cada perspectiva, los mismos que servirán para examinar y filtrar los indicadores (Abril Lara, 2016); estos son expuestos a los usuarios para que sean estos quienes determinen la relevancia o no de los datos encontrados en los OLTP. Además, se realiza una revisión de las preguntas obtenidas, ya que en ellas se aprecia los campos por los cuales se filtran los indicadores, así mismo los campos en el modelo relacional expuesto anteriormente, permitiendo la fácil deducción (Bustos Barrera & Mosquera Artieda, 2013).

5.8.2.4. Modelo conceptual ampliado

Se amplía el modelo conceptual generado con anterioridad colocando cada perspectiva y campos seleccionados, y bajo cada indicador la fórmula del cálculo mediante el cual se obtendrán dichos valores (Bernabeu, 2009) como se muestra en la Figura 25.

Figura 25

Estructura General de un Modelo Conceptual Ampliado



Nota. Adaptada de (Bernabeu, 2009).

5.8.3. Modelo lógico del Datawarehouse

El modelo lógico es una representación de la estructura de base de datos, que puede procesarse y almacenarse en algún Sistema Gestor de Base de Datos (*SGBD – Database Management*



System en inglés) (Silva Peñafiel, 2018). El modelo conceptual formulado en los pasos anteriores nos permitirá la construcción de este modelo lógico; para ello se selecciona el tipo de modelo lógico a ser utilizado, se diseñan las tablas tanto de dimensiones como de hechos y finalmente se establecen las uniones entre las tablas (Bustos Barrera & Mosquera Artieda, 2013).

5.8.3.1. Tipo del modelo lógico del DW

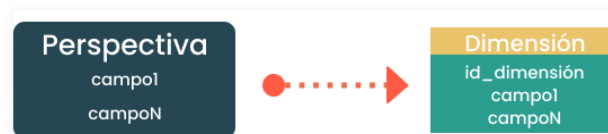
Los modelos lógicos permiten contener la estructura del Data Warehouse (Bernabeu, 2009). Hay que tener en cuenta que este modelo se ajuste a las necesidades y requerimientos de los usuarios. Es muy importante definir objetivamente si se empleará un esquema en estrella, constelación o copo de nieve, ya que esta decisión afectará considerablemente la elaboración del modelo lógico (Gonzales Farro, 2018).

- **Modelo lógico en estrella (star schema):** Habitualmente, se suele modelar la base de datos utilizando el esquema en estrella, en el que hay una única tabla central, la tabla de hechos, que contiene todas las medidas y una tabla adicional por cada una de las perspectivas desde las que queremos analizar dicha información, es decir por cada una de las dimensiones (Ramos, 2011).
- **Modelo lógico en constelación (star flake schema):** Este modelo está compuesto por una serie de esquemas en estrella, está formado por una tabla de hechos principal y por una o más tablas de hechos auxiliares, las cuales pueden ser sumalizaciones de la principal (Abril Lara, 2016). Dichas tablas se encuentran en el centro del modelo y están relacionadas con sus respectivas tablas de dimensiones.
- **Modelo lógico en copo de nieve (snowflake schema):** En este esquema las tablas de las dimensiones son normalizadas de tal manera que se simplifiquen las operaciones de selección de datos, es decir, se representa la información sin que exista redundancia de datos (Bustos Barrera & Mosquera Artieda, 2013). La diferencia es que algunas de las dimensiones no están relacionadas directamente con la tabla de hechos, sino que se relacionan con ella a través de otras dimensiones (Abril Lara, 2016). En este caso se tiene una tabla de hechos, situada en el centro, que contiene todas las medidas y una o varias tablas adicionales, con un mayor nivel de normalización (Bernabeu, 2009).

5.8.3.2. Tablas de Dimensiones

Se debe crear las dimensiones del esquema lógico tomando las perspectivas y campos definidos anteriormente de modo que (Bernabeu, 2009):

- Se procede a la elección de nombre que identifique la dimensión.
- Se integra un campo/atributo que represente la clave principal.
- Se redefinen los campos poco intuitivos de ser necesario.

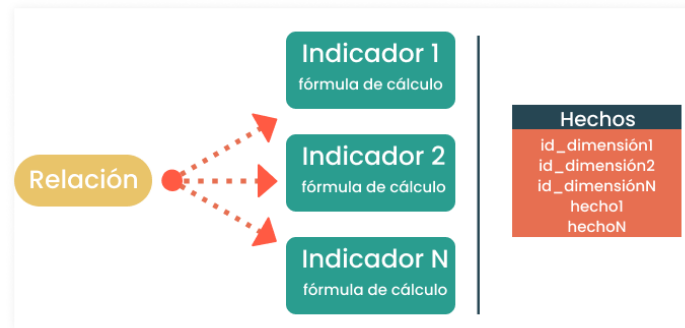
Figura 26*Modelo de Tabla de Dimensión*

Nota. Adaptada de (Bernabeu, 2009).

5.8.3.3. Tablas de Hechos

En esta etapa se diseñan las tablas de hechos que contienen los hechos a través de los cuales se construirán los indicadores de estudio (Bernabeu, 2009). Las tablas de hechos contienen, hechos que serán utilizados por los analistas para apoyar el proceso de toma de decisiones (Bustos Barrera & Mosquera Artieda, 2013). La definición de las tablas de hechos será determinada por el tipo de modelo lógico seleccionado, para el modelo en estrella es necesario (Silva Peñafiel, 2018):

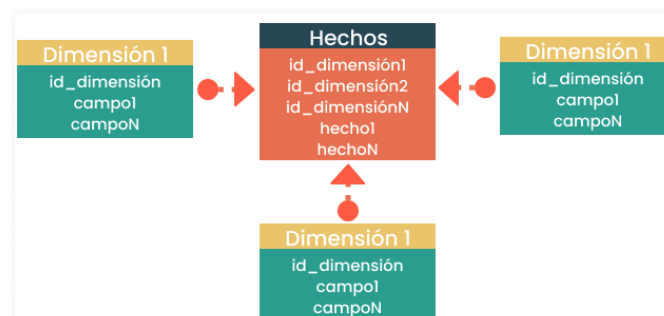
- Asignar un nombre a la tabla de hechos que indique la información analizada, área de estudio, etc.
- Definir una clave primaria, compuesta de la asociación de las claves primarias de cada tabla de dimensión relacionada.
- Crear campos como indicadores de acuerdo con los definidos en el modelo conceptual.

Figura 27*Modelo de Tabla de Hechos*

Nota. Adaptada de (Bernabeu, 2009).

5.8.3.4. Uniones

Este concepto se define como la relación correspondiente entre tablas de dimensiones y tablas de hechos (Abad León & Vizhñay Enderica, 2020). Este proceso se realiza para cualquier modelo lógico definido anteriormente y se puede resumir en la Figura 28.

Figura 28*Representación de Uniones para Tablas de Hechos y Dimensiones*

Nota. Adaptada de (Bernabeu, 2009).

5.8.4. Integración de datos

La carga de datos al modelo lógico anteriormente definido es el objetivo de este paso, es necesario emplear técnica de limpieza y calidad de datos, procesos ETL, etc. Así como las reglas y políticas para las continuas actualizaciones (Silva Peñafiel, 2018).



5.8.4.1. Carga inicial

Se deben establecer condiciones y restricciones para la carga de los datos hacia el Data Warehouse, se debe evitar la carga de datos con valores faltantes (missing values), anómalos (outliers) o sin integridad alguna (Bustos Barrera & Mosquera Artieda, 2013). Estos tipos de datos conllevan a una degradación de la calidad del DW en las etapas de consultas a las bases de datos y creación de reportes (Abad León & Vizhñay Enderica, 2020); el orden de la carga de estos datos debe darse de la siguiente manera: primero las tablas de dimensiones, continuando con las tablas de hechos, respetando las correspondencias entre cada elemento (Silva Peñafiel, 2018). Es de vital importancia describir paso a paso el proceso de carga de información hacia el Data Warehouse.

5.8.4.2. Actualización

Para la actualización de datos del Data Warehouse se debe procurar el establecimiento de políticas y estrategias (Silva Peñafiel, 2018); relacionadas con las actividades (Bernabeu, 2009):

- Especificar tareas de limpieza, calidad de datos, procesos ETL, etc.
- Especificación general y detallada de las acciones que cada software empleado para la carga de datos debe realizar.
- Establecer las políticas y estrategias a ejecutar como actualizaciones o manejo de datos (Bustos Barrera & Mosquera Artieda, 2013).

5.9. Minería de Datos y Obtención del Conocimiento

La Minería de Datos (*DM – Data Mining en inglés*) es un proceso no trivial de extracción de información que está oculta, previamente desconocida y es potencialmente útil, de grandes bases de datos (Agarwal, 2013). Por otra parte, (Hand & Adams, 2015) afirman que la minería de datos se encarga de analizar grandes conjuntos de datos, a menudo por la calidad deficiente de los datos, y por la amplitud del tipo de estructuras buscadas. (Uvidia Fassler, 2016) sostiene que la minería de datos proporciona beneficios evidentes para cualquier organización, ya que es un proceso de descubrimiento, que permite a los usuarios comprender la importancia de los datos, así como la relación existente entre los mismos, descubriendo patrones y rasgos en el contenido de esta información.

A partir de estos conceptos se puede deducir que Data Mining permite analizar factores de influencia en determinados procesos, predecir o estimar variables o comportamientos futuros, segmentar o agrupar ítems similares (Agarwal, 2013), además de obtener secuencias de eventos que



provocan comportamientos específicos, de donde se puede inferir modelos, relaciones y estimaciones de datos, para poder desarrollar predicciones sobre los mismos, sin la necesidad de contar con patrones o reglas preestablecidas y permitiendo tomar decisiones basadas en un conocimiento a partir de la información (Uvidia Fassler, 2016). Para (Chung & Gray, 1999) las etapas que incluye la minería de datos consisten en:

- a) Desarrollar una comprensión de la aplicación, del conocimiento previo relevante y de los objetivos del usuario final.
- b) Crear un conjunto de datos de destino que se utilizará para el descubrimiento.
- c) Limpiar y pre procesar datos (incluido el manejo de campos de datos faltantes, ruido en los datos, contabilidad de series de tiempo y cambios conocidos).
- d) Reducir el número de variables y encontrar representaciones invariantes de datos si es posible
- e) Elija la tarea de minería de datos (clasificación, regresión, agrupamiento, etc.)
- f) Elija el algoritmo de minería de datos.
- g) Buscar patrones de interés.
- h) Interpretar el patrón extraído

Los dos primeros pasos, comprender la aplicación y determinar los datos de destino que se utilizarán para el descubrimiento, son comunes a todas las formas de análisis de datos. Estos pasos incluyen saber qué conocimiento previo relevante existe, los objetivos del proyecto y qué datos existen (Chung & Gray, 1999). El tercer paso: (limpiar y pre procesar los datos) consiste en preparar los datos para su almacenamiento previo a un proceso de limpieza y validación, eliminando valores atípicos y ruido en los datos (Hand & Adams, 2015). Los pasos cuarto y quinto implican reducir el número de variables ya que, con demasiadas variables, resulta difícil dar sentido a los resultados. Por lo tanto, se intenta encontrar representaciones de los datos que sean invariantes (Agarwal, 2013). Los pasos sexto y séptimo (elegir el algoritmo de minería de datos y buscar patrones) es la parte esencial de la minería de datos. El algoritmo seleccionado depende de la tarea a realizar (Bernabeu, 2009). Implican ajustar modelos o determinar patrones a partir de los datos del almacén. En el octavo paso se examinan los resultados de la extracción de datos del paso anterior, se interpretan e informan los resultados. Estos pasos implican presentar los resultados en una forma que sea más comprensible quienes deben tomar decisiones basadas en información.



5.9.1. *Machine Learning*

Durante las últimas dos décadas, el aprendizaje automático se ha convertido en uno de los pilares de la tecnología de la información (Smola, 2008). Machine learning se puede definir ampliamente como métodos computacionales que utilizan la experiencia para mejorar el rendimiento o realizar predicciones precisas (Zhang, 2020). En el aprendizaje automático, generalmente se proporciona un conjunto de entrenamiento y un conjunto de prueba. Por conjunto de entrenamiento, significará la unión del conjunto etiquetado y el conjunto de ejemplos sin etiquetar disponibles. En comparación, el conjunto de prueba consta de ejemplos nunca vistos. El aprendizaje automático tiene como objetivo establecer un regresor o clasificador a través del aprendizaje del conjunto de entrenamiento y luego evaluar su rendimiento a través del conjunto de prueba (Mohri, Rostamizadeh, & Talwalkar, 2018).

Para datos estructurados (datos que a menudo se integran en bases de datos, se presentan de manera ordenada y poseen un formato) como los usados en este estudio, se puede clasificar el aprendizaje automático de la siguiente manera (Zhang, 2020).

- **Aprendizaje supervisado:** dado un conjunto de entrenamiento que consta de datos etiquetados (es decir, entradas de ejemplo y sus salidas deseadas), el aprendizaje supervisado aprende una regla general que asigna entradas a salidas.
- **Aprendizaje no supervisado:** en el aprendizaje no supervisado, el conjunto de entrenamiento consta del conjunto sin etiquetar. La tarea principal del aprendizaje automático es encontrar las soluciones por sí mismo (es decir, patrones, estructuras o conocimiento en datos sin etiquetar).
- **Aprendizaje semi supervisado:** Dado un conjunto de entrenamiento con una pequeña cantidad de datos etiquetados junto con una gran cantidad de datos no etiquetados. El aprendizaje semi supervisado se encuentra entre el aprendizaje no supervisado (sin ningún dato de entrenamiento etiquetado) y el aprendizaje supervisado (con datos de entrenamiento completamente etiquetados).
- **Aprendizaje reforzado:** los datos de entrenamiento (en forma de recompensas y castigos) se proporcionan sólo como retroalimentación a un agente de inteligencia artificial en un entorno dinámico. Esta retroalimentación entre el sistema de aprendizaje y la experiencia de interacción es útil para mejorar el desempeño en la tarea que se está aprendiendo. Q Learning es un popular aprendizaje por refuerzo sin modelos y aprende una función de recompensa o castigo (funciones de valor de acción, simplemente llamadas función Q).

- **Aprendizaje de Transferencia:** en muchas aplicaciones del mundo real, la distribución de datos cambia o los datos están desactualizados y, por lo tanto, es necesario aplicar el aprendizaje de transferencia para considerar la transferencia de conocimiento del dominio de origen al dominio de destino.

Cada uno de los tipos de aprendizaje antes mencionados poseen algoritmos específicos para su aplicación. Sin embargo, se hablará únicamente del aprendizaje no supervisado puesto que se han aplicado algoritmos de este tipo de aprendizaje en el estudio. Como ya se mencionó, este aprendizaje parte de un conjunto de datos del que no se tiene un conocimiento a priori, siendo el objetivo en este tipo de análisis la comprensión de los datos o la transformación automática de éstos (Roman, 2019). En este tipo de aprendizaje (Zhang, 2020) distingue las siguientes sub categorías de algoritmos no supervisados: clustering (dando una agrupación de los datos según su similitud), reducción de dimensionalidad (simplificando la estructura de los datos manteniendo sus características fundamentales), visualización (aplicación derivada de la reducción de dimensionalidad cuyo objetivo es mostrar gráficas de la estructura de un conjunto de datos multidimensional) y aprendizaje topológico (extrayendo la estructura interna con la que se distribuyen los datos en su espacio original).

Para el estudio se ha hecho de técnicas de clustering. El clustering (agrupación), es un proceso de agrupación de datos de entrada en grupos de similares características. Esta técnica, no supervisada, no necesita ningún atributo que caracteriza la clase de equivalencia a la que pertenece cada una de las instancias entrantes, puesto que su salida es definir este atributo (Roman, 2019). Para ello, cada clúster (o grupo) está formado por una colección de objetos que son similares (o se consideran similares) entre sí, pero que son distintos respecto a los objetos de clústeres diferentes.

5.9.1.1. Algoritmo K-Means

Uno de los algoritmos más relevantes en clustering es K-Medias (*K-Means en inglés*). El algoritmo K-Means tiene como objetivo encontrar y agrupar en clases los puntos de datos que tienen una alta similitud entre ellos. En términos del algoritmo, esta similitud se entiende como lo opuesto de la distancia entre puntos de datos. Cuanto más cerca estén los puntos de datos, más similares y con más probabilidades de pertenecer al mismo clúster serán (Smola, 2008). Este algoritmo se utiliza en conjuntos de datos no etiquetados, es decir, datos sin categorías o grupos definidos. El objetivo de este algoritmo es encontrar grupos en los datos, con el número de grupos representados por la variable k . Para utilizar K-Means se especifica el número de grupos a encontrar. De manera general, el algoritmo contempla las siguientes etapas (Zhang, 2020).



- **Inicialización:** se elige la localización de los centroides (nombre que reciben los centros de los clusters según el criterio de distancia aplicado) de los k grupos aleatoriamente.
- **Asignación:** se asigna cada dato al centroide más cercano.
- **Actualización:** se actualiza la posición del centroide a la media aritmética de las posiciones de los datos asignados al grupo.

Las etapas de Asignación y Actualización se siguen iterativamente hasta que no haya más cambios. En otras palabras, este algoritmo concluye en cuanto los clusters son estables y no se han modificado los puntos medios a lo largo de dos interacciones. Los resultados de este algoritmo son (Romero, Ventura, & García, 2008):

- Los centroides de los clústeres k , que pueden ser usados para etiquetar nuevos datos.
- Etiquetas para los datos de formación, cada punto de datos se asigna a un único clúster.

Cada centroide de un clúster es un conjunto de valores de características que definen los grupos resultantes. Examinar las ponderaciones de las características del centroide puede utilizarse para interpretar cualitativamente qué tipo de grupo representa cada conglomerado (G. de Barba, y otros, 2020). El algoritmo K-means encuentra los clústeres y las etiquetas de conjuntos de datos para un k en particular preseleccionado. En general, no existe un método para determinar el valor exacto de k , pero se puede obtener una estimación precisa con ayuda de algunas técnicas.

Una de las métricas que se utilizan comúnmente para comparar diferentes valores de k es la distancia media entre los puntos de datos y el centroide de su clúster. Esta métrica no puede utilizarse como único objetivo. En su lugar, se traza la distancia media al centroide en función de k y se puede utilizar el método del codo para determinar aproximadamente k . Existen otras técnicas como la validación cruzada, criterios de información, el método de salto teórico de información y el método de la silueta (Luna, Castro, & Romero, 2017).

Las desventajas de este algoritmo son (Mohri, Rostamizadeh, & Talwalkar, 2018):

- Es necesario definir el número de clústeres. Esta decisión puede afectar a los resultados.
- Como la ubicación inicial de los centroides es aleatoria, los resultados pueden no ser comparables y mostrar una falta de consistencia.

- K-means produce clústeres con tamaños uniformes, cada clúster tiene aproximadamente la misma cantidad de observaciones, aunque los datos pueden comportarse de manera diferente, y es muy sensible a valores atípicos y datos ruidosos.

5.9.1.2. Técnicas de validación de modelos de agrupación

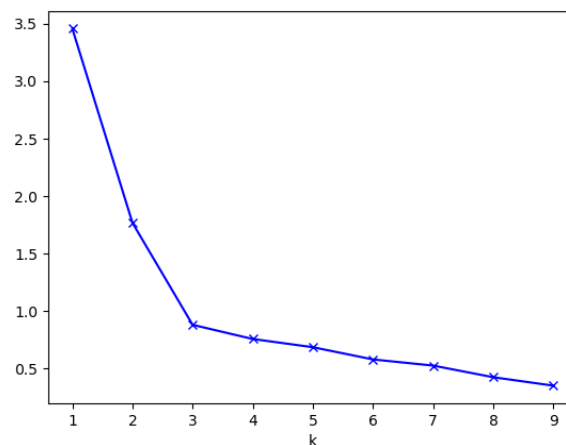
El Clustering (agrupamiento) es especialmente sensible a los parámetros de la entrada (Štrobl, Piorecký, & Krajča, 2017). Es de importancia evaluar el resultado de los algoritmos de clustering, sin embargo, es difícil definir cuando el resultado de un agrupamiento es aceptable. Por esta razón existen técnicas e índices para la validación de un agrupamiento realizado (Shi, y otros, 2021).

Elbow Curve (Método del Codo)

También denominado “Método del Codo”; probablemente es el más conocido. Este método calcula y grafica la suma de cuadrado en cada número de clústeres, en este gráfico se busca un cambio de pendiente de empinada a poca profundidad (un codo) para determinar el número óptimo de clústeres (Et-taleby, Boussetta, & Benslimane, 2020). Este método es inexacto, pero sigue siendo potencialmente útil. Este método muestra cómo el aumento del número de clústeres contribuye a separar los clústeres de una manera significativa (Shi, y otros, 2021). Como se puede observar en la Figura 29, la curva indica que los grupos adicionales más allá del tercero tienen poco valor.

Figura 29

Elbow curve, para el cual en $k = 3$ será el número de clusters óptimo.



Nota. Tomada de (Štrobl, Piorecký, & Krajča, 2017).

Silhouette Curve (Método de la Silueta)

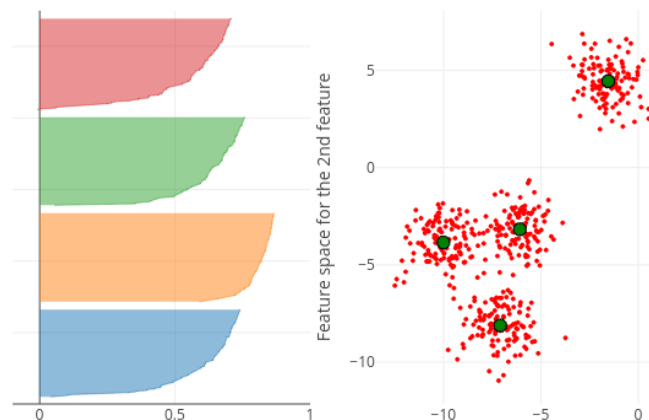
El método de la silueta puede utilizarse para estudiar la distancia de separación entre los grupos resultantes. El gráfico de silueta muestra una medida de cuán cerca está cada punto de un clúster de los puntos en los clústeres vecinos y por lo tanto proporciona una forma de evaluar parámetros como el número de clústeres visualmente (Marutho, Hendra Handaka, Wijaya, & Muljono, 2018). Esta medida varía de -1 a 1 . Donde 1 significa que los grupos están muy cerca de su propio clúster y lejos de otros, mientras -1 indica que los puntos están cerca de los clústeres vecinos (Štrobl, Piorecký, & Krajča, 2017). Se considera que este método es mejor, ya que hace que la decisión sobre el número óptimo de clústeres sea más significativa y clara. El coeficiente de silueta se calcula utilizando la distancia media intra-clúster a y la distancia media más cercana al clúster b para cada muestra. El coeficiente de silueta para una muestra es (Et-taleby, Boussetta, & Benslimane, 2020):

$$\text{coeficiente} = \frac{(b - a)}{\max(a, b)}$$

Donde b es la distancia entre una muestra y el clúster más cercano del que la muestra no forma parte. Este coeficiente se calcula para todas las muestras y se usa como una métrica para juzgar el número de grupos (Shi, y otros, 2021). Como se puede observar en la Figura 32, para $k = 4$ la distribución de grupos es uniforme y el valor del coeficiente es cercano a la unidad por lo que se puede determinar que el número óptimo de clústers es $k = 4$.

Figura 30

Silhouette Curve, para el cual en $k = 4$ será el número de clusters óptimo.



Nota. Adaptada de (Shi, y otros, 2021).



Capítulo 6: Desarrollo de un Dashboard de Visualización

6.1. Análisis Exploratorio de los Datos

Se ha hecho uso de las dos primeras etapas de la metodología Hefesto para el análisis exploratorio de la base de datos de Moodle. La etapa de [Análisis de Requerimientos](#) y la etapa de [Análisis de los OLTP](#), con el objetivo de comprender la información, su origen, sus relaciones y demás aspectos relevantes para la construcción del dashboard de visualización. Para el análisis exploratorio se contó con una base de datos de Moodle que almacena datos del curso virtual Planificación y evaluación de clases virtuales (2020-05) de la plataforma Edu-Analytics¹. Si bien los datos analizados y las conclusiones encontradas se basan en la estructura de base de datos de la plataforma Edu-Analytics, la información detallada en este documento es fácilmente aplicable a cualquier base de datos de Moodle.

El sitio Moodle de Edu-Analytics utiliza PostgreSQL como gestor de Base de Datos. Se proporcionó una copia de la base de datos luego de haber finalizado el curso Planificación y evaluación de clases virtuales (2020-05), la base de datos cuenta con registros de 41 estudiantes matriculados al curso. El curso tuvo una duración de dos semanas. La base de datos contiene información relacionada a información del curso, participantes (docente y estudiantes), contenido del curso en base a [módulos/actividades](#) (tareas, libros, chat, elección, retroalimentación, folder, foro, retroalimentación, glosario, página, cuestionario/evaluación, recurso, scorm, encuesta, url, wiki, workshop), interacciones entre participantes con el contenido del curso y configuración del sitio. El curso involucró la mayor cantidad de módulos/actividades con el objetivo de mejorar el análisis.

6.1.1. Análisis de requerimientos

Inicialmente fue de gran importancia identificar los requerimientos de los usuarios a través de preguntas que reflejan los objetivos del presente trabajo. Posteriormente, se analizaron estas preguntas a fin de identificar indicadores y perspectivas. Finalmente se confeccionó un modelo conceptual en donde se visualiza el resultado obtenido en este primer paso.

6.1.1.1. Identificar preguntas

El objetivo principal de esta fase fue la de obtener e identificar las necesidades de información clave de alto nivel, que fueron esenciales para llevar a cabo los objetivos y responder las preguntas

¹ Edu-Analytics: <https://virtual.edu-analytics.ec/>



de investigación planteadas. Se formularon preguntas, que incluyen variables de análisis que se consideraron relevantes.

Para la formulación de las preguntas se ha considerado el modelo pedagógico de [clase invertida](#), además de la autorregulación por parte de los estudiantes. En base a esto, se hizo la siguiente clasificación:

- **Actividades de estudio AE:** indicadores relacionados a la actividad de los estudiantes en el curso en términos de visitas realizadas, tiempo invertido y sesiones a lo largo del curso en diferentes intervalos de tiempo.
- **Seguimiento (tiempo) ST:** indicadores relacionados al tiempo invertido en el curso en relación a un tiempo planificado por el docente en términos de semanas.
- **Seguimiento (actividades) SA:** indicadores relacionados a la entrega de actividades asignadas por parte del docente a los estudiantes del curso y acceso a recursos puestos a disposición de los estudiantes con el objetivo de evidenciar la participación de los estudiantes con el material didáctico del curso.
- **Seguimiento (evaluaciones) SE:** indicadores relacionados a los promedios de calificaciones basados en las actividades evaluables en Moodle, además de un resumen de intentos en las diferentes evaluaciones del curso y análisis de dichas preguntas de una evaluación.

Actividades de estudio

P_AE_1. Se requiere saber la cantidad de visitas realizadas por los estudiantes al curso. La información debe poder visualizarse por (día de la semana) / (hora del día) y por mes / (número de semana). También es necesario que se pueda visualizar el número de sesiones.

P_AE_2. Se requiere visualizar un reporte con el número de visitas al curso de los estudiantes, la cantidad de tiempo han empleado en el curso y el progreso en el curso basado en las actividades realizadas y recursos revisados.

P_AE_3. Se desea conocer el número de sesiones menores a 30 minutos, mayores a 30 y menores a 60 minutos y mayores a 60 minutos en un curso.

Seguimiento (tiempo)



P_ST_1. Se desea conocer el tiempo promedio invertido en el curso en comparación al tiempo promedio planificado por el docente a lo largo de las diferentes semanas planificadas en el curso.

Seguimiento (actividades)

P_SA_1. Se desea conocer la cantidad de estudiantes que han enviado a tiempo, tarde o que no han enviado actividades de tipo **assign** en cada una de las diferentes semanas planificadas a lo largo del curso. Además, permitir la posibilidad de enviar un email a los estudiantes que han enviado, no han enviado o han enviado tarde una actividad de tipo **assign**.

P_SA_2. Se desea conocer la cantidad de estudiantes que han accedido y no han accedido a los recursos y actividades (que el docente seleccione) durante las diferentes semanas de duración del curso. Además, permitir la posibilidad de enviar un email a los estudiantes que han accedido o no a dichos recursos y actividades.

Seguimiento (para evaluaciones)

P_SE_1. Se desea conocer el promedio general de calificaciones de cada una de las actividades evaluables del curso. Al seleccionar una actividad específica mostrar la distribución de calificaciones con el promedio, la nota máxima y la nota mínima para dicha actividad.

P_SE_2. Se desea conocer la cantidad de total de intentos, intentos parcialmente correctos, intentos incorrectos, intentos correctos, preguntas en blanco y sin calificar en cada una de las preguntas de una evaluación de todos los intentos realizados por los estudiantes.

6.1.1.2. Identificar indicadores y perspectivas

Con las preguntas planteadas, se procede a descomponer dichas preguntas para descubrir los indicadores que se utilizarán y las perspectivas de análisis.



indicadores



perspectivas



Actividades de estudio

P_AE_1. Se requiere saber la *cantidad de visitas realizadas* por los *estudiantes* al *curso*. La información debe poder visualizarse por (*día de la semana*) / (*hora del día*) y por mes / (*número de semana*). También es necesario que se pueda visualizar el *número de sesiones*.

Indicadores Cantidad de visitas realizadas
Número de sesiones

Perspectivas Curso
Estudiante
Tiempo

P_AE_2. Se requiere visualizar un reporte con el *número de visitas* al *curso* de los *estudiantes*, la *cantidad de tiempo han empleado en el curso*, y el *progreso en el curso* basado en las actividades realizadas y recursos revisados.

Indicadores Número de visitas
Cantidad de tiempo que han pasado en el curso
Progreso en el curso

Perspectivas Curso
Estudiante
Tiempo

P_AE_3. Se desea conocer el *número de sesiones menores a 30 minutos, mayores a 30 y menores a 60 minutos y mayores a 60 minutos* en un *curso*.

Indicadores Número de sesiones

Perspectivas Tipo de Sesión (< 30 min, > 30 min y < a 60 min, > 60 min)
Estudiante
Tiempo
Curso

Seguimiento (tiempo)

P_ST_1. Se desea conocer el *tiempo promedio invertido* en el *curso* en comparación al *tiempo promedio planificado* por el *docente* a lo largo de las diferentes *semanas* planificadas en el curso.



Indicadores Tiempo promedio invertido
Tiempo promedio planificado

Perspectivas Docente
Tiempo
Curso

Seguimiento (actividades)

P_SA_1. Se desea conocer la *cantidad de estudiantes* que han enviado a tiempo, tarde o que no han enviado *actividades de tipo assign* en cada una de las diferentes *semanas* planificadas a lo largo de todo el *curso*. Además, permitir la posibilidad de enviar un email a los estudiantes que han enviado, no han enviado o han enviado tarde una actividad de tipo *assign*.

Perspectivas Estudiante
Tiempo
Curso
Actividad de tipo assign
Tipo Envío (a tiempo/fuera de tiempo/sin envío)

P_SA_2. Se desea conocer la *cantidad de estudiantes* que han accedido y no han accedido a los *recursos y actividades* (que el docente seleccione) durante las diferentes *semanas* de duración del *curso*. Además, permitir la posibilidad de enviar un email a los estudiantes que han accedido o no a dichos recursos y actividades.

Perspectivas Estudiante
Tiempo
Curso
Recurso
Actividades
Tipo de acceso (Con Acceso/ Sin Acceso)

Seguimiento (para evaluaciones)

P_SE_1. Se desea conocer el *promedio general de calificaciones* de cada una de las *actividades* evaluables del *curso*. Al seleccionar una actividad específica mostrar la *distribución de calificaciones* con el promedio, la nota máxima y la nota mínima para dicha actividad.



Indicadores Promedio general de calificaciones en actividades evaluables
Distribución de calificaciones

Perspectivas Actividad
Estudiante
Curso

P_SE_2. Se desea conocer la *cantidad de total de intentos, intentos parcialmente correctos, intentos incorrectos, intentos correctos, preguntas en blanco y sin calificar* en cada una de las *preguntas* de una *evaluación* de todos los intentos realizados por los *estudiantes*.

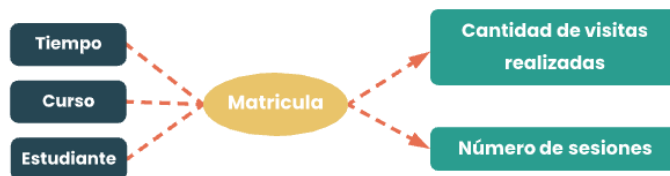
Perspectivas Tipo de Intento (correcto/parcialmente correcto/incorrecto/en blanco/ sin calificar)
Pregunta
Evaluación
Estudiante
Curso

6.1.1.3. Modelo conceptual

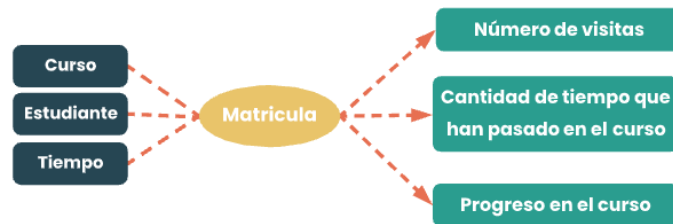
En esta etapa, se construirá un modelo conceptual a partir de los indicadores y perspectivas obtenidas en el paso anterior.

Actividades de estudio

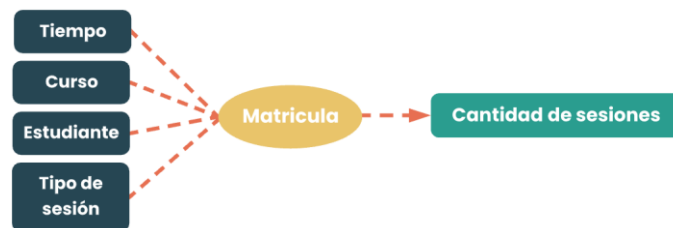
P_AE_1. Se requiere saber la *cantidad de visitas realizadas* por los *estudiantes* al *curso*. La información debe poder visualizarse por *(día de la semana) / (hora del día) y por mes / (número de semana)*. También es necesario que se pueda visualizar el *número de sesiones*.



P_AE_2. Se requiere visualizar un reporte con el *número de visitas* al *curso* de los *estudiantes*, la *cantidad de tiempo han empleado en el curso*, y el *progreso en el curso* basado en las actividades realizadas y recursos revisados.

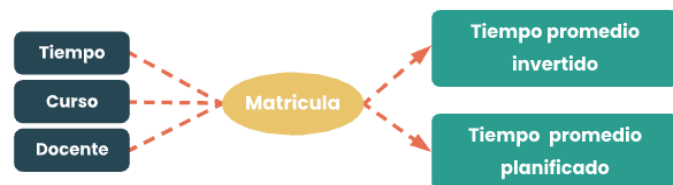


P_AE_3. Se desea conocer el *número de sesiones menores a 30 minutos, mayores a 30 y menores a 60 minutos y mayores a 60 minutos* en un *curso*.



Seguimiento (tiempo)

P_ST_1. Se desea conocer el *tiempo promedio invertido* en el *curso* en comparación al *tiempo promedio planificado* por el *docente* a lo largo de las diferentes *semanas* planificadas en el curso.



Seguimiento (actividades)

P_SA_1. Se desea conocer la *cantidad de estudiantes* que han enviado a tiempo, tarde o que no han enviado *actividades de tipo assign* en cada una de las diferentes *semanas* planificadas a lo



largo de todo el **curso**. Además, permitir la posibilidad de enviar un email a los estudiantes que han enviado, no han enviado o han enviado tarde una actividad de tipo **assign**.

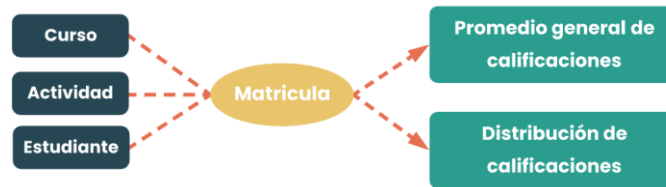


P_SA_2. Se desea conocer la **cantidad de estudiantes** que han accedido y no han accedido a los **recursos y actividades** (que el docente seleccione) durante las diferentes **semanas** de duración del **curso**. Además, permitir la posibilidad de enviar un email a los estudiantes que han accedido o no a dichos recursos y actividades.



Seguimiento (para evaluaciones)

P_SE_1. Se desea conocer el **promedio general de calificaciones** de cada una de las **actividades** evaluables del **curso**. Al seleccionar una actividad específica mostrar la **distribución de calificaciones** con el promedio, la nota máxima y la nota mínima para dicha actividad.



P_SE_2. Se desea conocer la *cantidad de total de intentos, intentos parcialmente correctos, intentos incorrectos, intentos correctos, preguntas en blanco y sin calificar* en cada una de las *preguntas* de una *evaluación* de todos los intentos realizados por los *estudiantes*.



6.1.2. Análisis de los OLTP

Seguidamente, se analizarán las fuentes OLTP para determinar cómo serán calculados los indicadores y para establecer las respectivas correspondencias entre el modelo conceptual creado en el paso anterior y las fuentes de datos. Luego, se definirán qué campos se incluirán en cada perspectiva. Finalmente, se ampliará el modelo conceptual con la información obtenida en este paso.

Para el análisis de las interacciones en la base de datos de Moodle, se realizó un análisis de la tabla que almacena las interacciones que se producen en el curso:

mdl_logstore_standard_log

Esta tabla almacena los registros de cada una de las interacciones que se producen en el sitio Moodle. Los campos importantes de esta tabla se muestran en la Tabla 6.

**Tabla 6**

Descripción de los Campos más Relevantes de la tabla mdl_logstore_standard_log

Campo / Columna	Descripción
id	Identificador en la tabla para cada evento.
eventname	Nombre del evento desencadenado en el curso. Este campo se encuentra conformado por las columnas component , target y action de la siguiente manera: <code>\component\event\target_action</code>
component	Componente con el que se hizo la interacción, algunos ejemplos son: mod_quiz, mod_forum, mod_url, mod_resource, mod_page, mod_scomm, mod_folder, mod_assign, core, report_participation, gradereport_overview, report_outline, assignsubmission_file, gradereport_user, etc. Representa las acciones de interacción sobre los componentes (esto se explica más a detalle en la siguiente sección), algunos ejemplos son:
action	assigned, updated, exported, deleted, submitted, restored, abandoned, viewed, shown, uploaded, unassigned, ended, graded, started, created, graph, searched, downloaded, reviewed Representa el componente de moodle que fue el objeto de la interacción, estos pueden ser:
target	block_analytics_graphs_event_view, user_list, remove_submission_form, all_files, course_user_report, role, report, attempt, feedback, grading_table, discussion_subscription, comment, badge_listing, post, user_profile, user, submission_form, grade, calendar_event, step, course_section, all_submissions, question, enrol_instance, edit_page, course_resources_list, grade_report, course_module_instance_list, assessable, course, grade_item, activity_report, tour, question_manually, course_backup, subscription, submission, course_bin_item, grading_form, attempt_preview, course_module, user_enrolment, submission_status, attempt_summary, discussion Representa la tabla que almacena el registro específico del componente objeto de la interacción, en otras palabras, son las tablas de la base de datos de Moodle con el prefijo mdl_, estas pueden ser:
objecttable	url, quiz_attempts, resource, assign_grades, grade_grades, page, role, comments, enrol, grade_items, forum_subscriptions, tool_usertours_tours, course_sections, forum, assign_submission, assignsubmission_file, user, user_enrolments, forum_discussions, tool_usertours_steps, forum_discussion_subs, forum_posts, assign, question, event, course, course_modules, quiz, tool_recyclebin_course, folder
objectid	Identificador de la instancia correspondiente de la tabla objeto de la interacción (por ejemplo, si objecttable es assign, objectid es el id del registro correspondiente en la tabla mdl_assign)
courseid	Identificador del curso donde se realiza la interacción
userid	Identificador del usuario que realizó la interacción.
relateduserid	Identificador del usuario relacionado a la interacción (por ejemplo, si la interacción es una vista de un perfil, relateduserid almacena el identificador del usuario de quien se vio el perfil).
timecreated	Fecha de creación del registro de interacción.



Posteriormente, se identificaron las interacciones con la mayoría de módulos/actividades en un curso. Estas interacciones difieren del usuario que las realiza, para nuestro estudio estudiante y docente. De manera general, las interacciones que se almacenan son las siguientes:

Estudiantes

- *Ingresos* al curso de moodle (loggedin).
- Interacciones con *actividades tipo assign* (ver/enviar/comentar tarea, ver retroalimentación, cargar archivos de una tarea).
- Interacciones con el *curso* (ver el curso, ver lista de usuarios).
- Interacciones con *actividades tipo folder* (ver/descargar folder).
- Interacciones con *actividades tipo foro* (ver foro, crear/borrar suscripción a un foro o una discusión de foro, ver/crear/eliminar discusión en un foro, crear/editar/eliminar publicación en un foro).
- Calificaciones entre estudiantes (*actividades tipo workshop*).
- Interacciones con *actividades tipo page* (ver page).
- Interacciones con *actividades tipo quiz* (ver quiz, realizar un intento a un quiz, ver página de resumen de un quiz, enviar un quiz).
- Interacciones con *actividades tipo resource* (ver resource).
- Interacciones con *tour del curso* (iniciar un tour, dar un paso en el tour, finalizar tour).
- Interacciones con *actividades tipo url* (ver url).
- Interacciones con *usuarios* (ver perfil de un usuario).
- Interacciones generales (ver reportes de calificación, ver reporte de usuarios del curso, ver estado de tareas, etc).

Docentes

- *Ingresos* a Moodle (loggedin).
- Interacciones con *actividades tipo assign* (ver/descargar/calificar tarea).
- Interacciones con *módulos de un curso* (crear/actualizar un módulo del curso).
- Interacciones con *secciones de un curso* (crear/actualizar/eliminar sección).
- Interacciones con *eventos de calendario de un curso* (crear/actualizar evento).
- Interacciones con *actividades tipo folder* (ver folder, descargar folder).



- Interacciones con *actividades tipo foro* (ver foro, crear/borrar suscripción a un foro o una discusión de foro, ver/crear/eliminar discusión en un foro, crear/editar/eliminar publicación en un foro).
- Interacciones con *ítems evaluables de estudiantes* (crear/eliminar calificación).
- Interacciones con *actividades tipo page* (ver page).
- Interacciones con *actividades tipo quiz* (ver/revisar/eliminar un intento de quiz, realizar un intento a un quiz (modo previsualización), ver página de resumen de un quiz (modo previsualización)).
- Interacciones con *preguntas de un quiz* (ver/crear/actualizar/calificar pregunta).
- Interacciones con *actividad tipo resource* (ver resource).
- Interacciones con *asignación de roles* (asignar rol a un usuario).
- Interacciones con *actividades tipo url* (ver url).
- Interacciones con *usuarios* (ver perfil de un usuario, matricular usuario).
- Interacciones generales (ver página de edición de módulos, ver/generar reportes, ver estados de envío de tareas, etc).

Las interacciones que se registran en la tabla `mdl_logstore_standard_log` están limitadas a la configuración del curso Planificación y evaluación de clases virtuales (2020-05). Diferentes configuraciones del curso o de contenidos del curso podrían permitir registrar mayor o menor cantidad de interacciones. Sin embargo, las interacciones analizadas fueron relevantes para responder las preguntas de la sección [Análisis de requerimientos](#).

Gracias al análisis realizado a la tabla de interacciones y sus relaciones con otras tablas de la base de datos de Moodle, se logró identificar las tablas de datos relevantes para el estudio, de esta manera se descartaron un gran número de tablas que no guardaban relación con el estudio. Para una mejor comprensión, se aislaron estas tablas de interés, se resumieron y destacaron sus atributos más importantes (véase Anexo 3).

Realizado todo el análisis, relaciones y resumen de atributos relevantes de las tablas de datos ya se tiene una comprensión a gran escala de cómo funcionan e interactúan los datos en la base de datos de Moodle. A partir de todo este proceso previo se continuó con la subetapas de la metodología Hefesto.



6.1.2.1. Conformar indicadores

Actividades de Estudio

P_AE_1. Se requiere saber la *cantidad de visitas realizadas* por los *estudiantes* al *curso*. La información debe poder visualizarse por *(día de la semana) / (hora del día) y por mes / (número de semana)*. También es necesario que se pueda visualizar el *número de sesiones*.

“Cantidad de visitas realizadas”

Hecho: Cantidad de visitas realizadas

Función de sumarización: SUM.

Aclaración: representa la sumatoria de las visitas realizadas al curso por parte de los estudiantes de un curso Moodle.

“Número de sesiones realizadas”

Hecho: Número de sesiones

Función de sumarización: SUM.

Aclaración: representa la sumatoria de sesiones realizadas al curso por parte de los estudiantes de un curso de manera general y para un estudiante específico.

P_AE_2. Se requiere visualizar un reporte con el *número de visitas* al *curso* de los *estudiantes*, la *cantidad de tiempo han empleado en el curso*, y el *progreso en el curso* basado en las actividades realizadas y recursos revisados.

“Número de visitas realizadas”

Hecho: Número de visitas

Función de sumarización: SUM.

Aclaración: representa la sumatoria de visitas realizadas al curso por parte de cada uno de los estudiantes de un curso.

“Cantidad de tiempo empleado en el curso”

Hecho: Cantidad de tiempo empleado en el curso

Función de sumarización: SUM.

Aclaración: representa la sumatoria de tiempo empleado por parte de cada uno de los estudiantes desde que inició el curso.

“Progreso en el curso”

Hecho: Progreso en el curso

Función de sumarización: SUM.

Aclaración: representa el progreso de un estudiante en base a las actividades realizadas en el curso.



P_AE_3. Se desea conocer el *número de sesiones menores a 30 minutos, mayores a 30 y menores a 60 minutos y mayores a 60 minutos* en un *curso*.

“Cantidad de Sesiones”

Hecho: Cantidad de Sesiones

Función de sumarización: COUNT

Aclaración: representa la cantidad de sesiones de los estudiantes en un curso. Estas sesiones son de tres tipos: menores a 30 minutos, mayores a 30 minutos y mayores a 60 minutos.

Seguimiento (tiempo)

P_ST_1. Se desea conocer el *tiempo promedio invertido* en el *curso* en comparación al *tiempo promedio planificado* por el *docente* a lo largo de las diferentes *semanas* planificadas en el curso.

“Tiempo promedio invertido”

Hecho: Tiempo promedio invertido

Función de sumarización: AVG.

Aclaración: representa la cantidad promedio de tiempo invertido por parte de los estudiantes en términos del tiempo planificado por el docente.

“Tiempo promedio planificado”

Hecho: Tiempo promedio planificado

Función de sumarización: AVG.

Aclaración: representa el tiempo promedio que el docente calcula en su planificación y que espera que los estudiantes inviertan durante su planificación.

Seguimiento (actividades)

P_SA_1. Se desea conocer la *cantidad de estudiantes* que han enviado a tiempo, tarde o que no han enviado *actividades de tipo assign* en cada una de las diferentes *semanas* planificadas a lo largo de todo el *curso*. Además, permitir la posibilidad de enviar un email a los estudiantes que han enviado, no han enviado o han enviado tarde una actividad de tipo *assign*.

Hecho: para responder a esta pregunta no se ha visto necesario crear hechos

Función de sumarización: COUNT.



P_SA_2. Se desea conocer la *cantidad de estudiantes* que han accedido y no han accedido a los *recursos y actividades* (que el docente seleccione) durante las diferentes *semanas* de duración del *curso*. Además, permitir la posibilidad de enviar un email a los estudiantes que han accedido o no a dichos recursos y actividades.

Hecho: para responder a esta pregunta no se ha visto necesario crear hechos

Función de sumarización: COUNT.

Seguimiento (para evaluaciones)

P_SE_1. Se desea conocer el *promedio general de calificaciones* de cada una de las *actividades* evaluables del *curso*. Al seleccionar una actividad específica mostrar la *distribución de calificaciones* con el promedio, la nota máxima y la nota mínima para dicha actividad.

“Promedio general de calificaciones”

Hecho: Promedio general de calificaciones

Función de sumarización: AVG.

Aclaración: representa el promedio general de calificaciones de para las diferentes tareas o cuestionarios (evaluaciones) del curso.

“Distribución de calificaciones”

Hecho: Distribución de calificaciones

Aclaración: representa la nota máxima, la nota mínima y el promedio de calificación en de alguna de las actividades evaluables.

P_SE_2. Se desea conocer la *cantidad de total de intentos, intentos parcialmente correctos, intentos incorrectos, intentos correctos, preguntas en blanco y sin calificar* en cada una de las *preguntas* de una *evaluación* de todos los intentos realizados por los *estudiantes*.

Hecho: para responder a esta pregunta no se ha visto necesario crear hechos

Función de sumarización: COUNT.

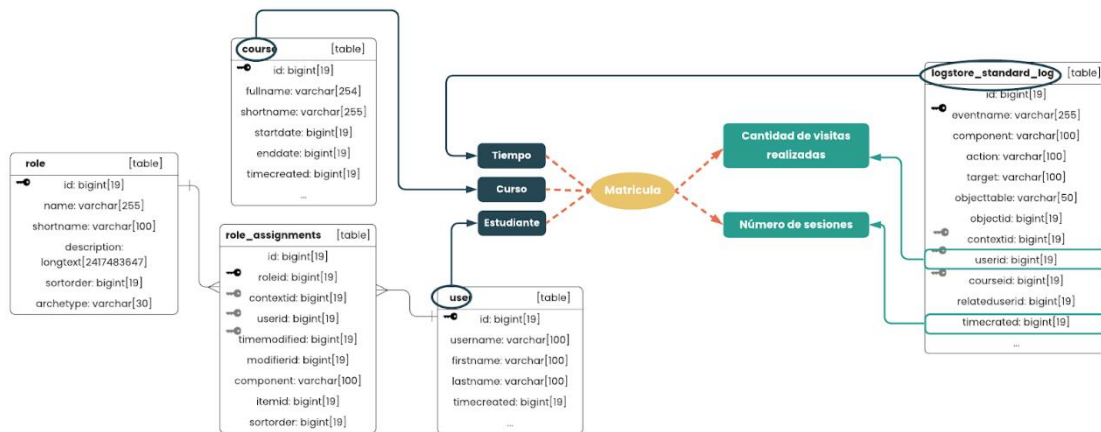
6.1.2.2. Establecer correspondencias

El objetivo de este paso, es el de examinar los OLTP disponibles que contengan la información requerida, como así también sus características, para poder identificar las correspondencias entre el modelo conceptual y las fuentes de datos. La idea es, que todos los elementos del modelo conceptual estén correspondidos en los OLTP.

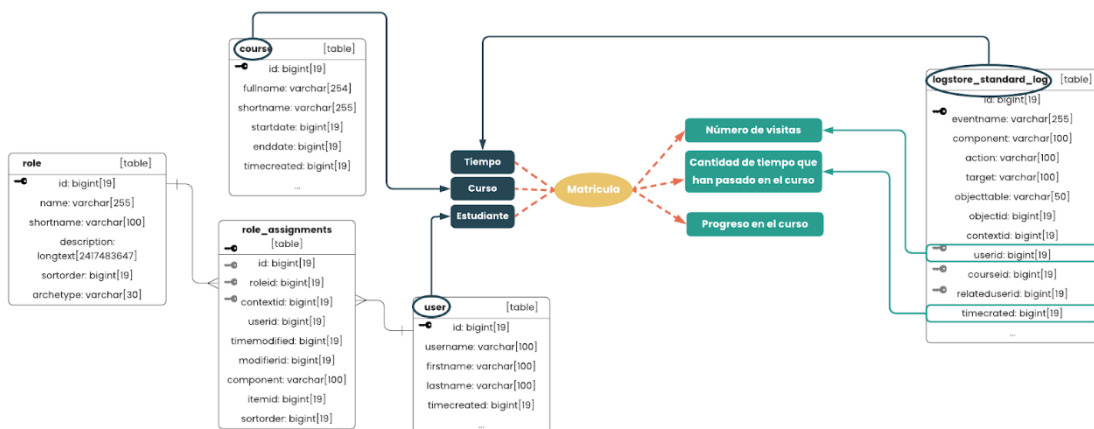


Actividades de estudio

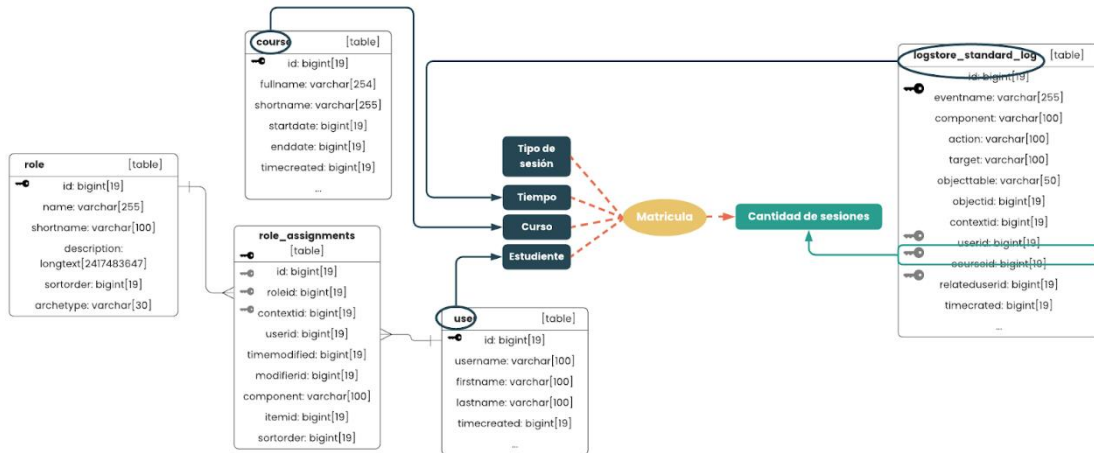
P_AE_1. Se requiere saber la *cantidad de visitas realizadas* por los *estudiantes* al *curso*. La información debe poder visualizarse por (*día de la semana*) / (*hora del día*) y por mes / (*número de semana*). También es necesario que se pueda visualizar el *número de sesiones*.



P_AE_2. Se requiere visualizar un reporte con el *número de visitas* al *curso* de los *estudiantes*, la *cantidad de tiempo han empleado en el curso*, y el *progreso en el curso* basado en las actividades realizadas y recursos revisados.

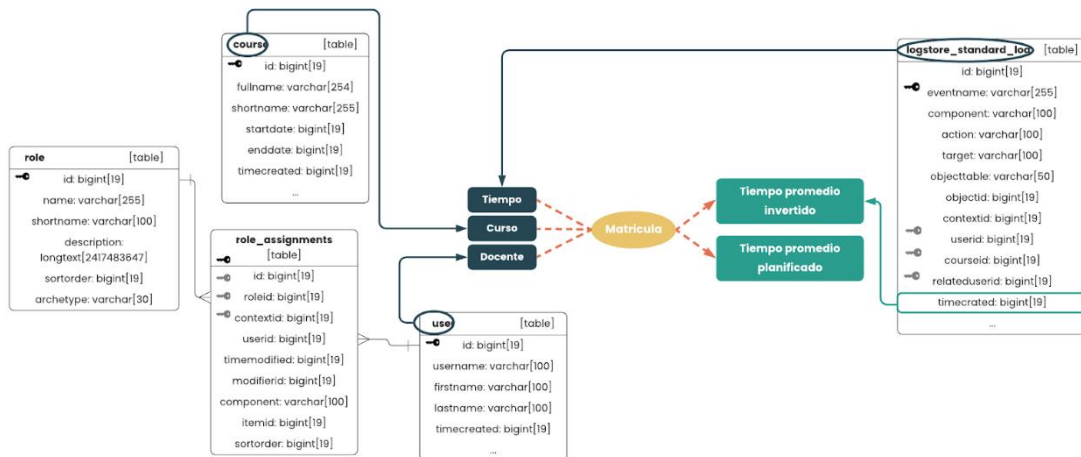


P_AE_3. Se desea conocer el *número de sesiones menores a 30 minutos, mayores a 30 y menores a 60 minutos y mayores a 60 minutos* en un *curso*.



Seguimiento (tiempo)

P_ST_1. Se desea conocer el *tiempo promedio invertido* en el *curso* en comparación al *tiempo promedio planificado* por el *docente* a lo largo de las diferentes *semanas* planificadas en el curso.

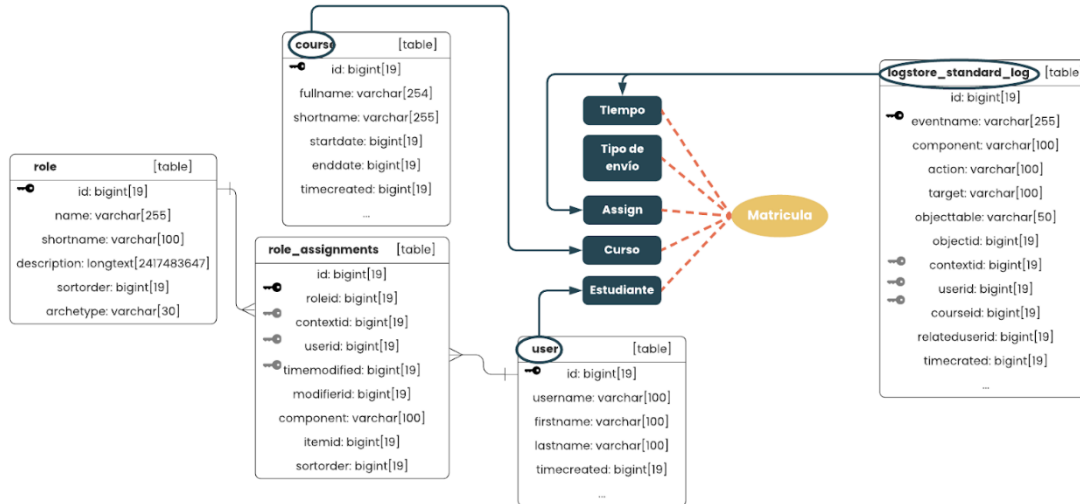


Seguimiento (actividades)

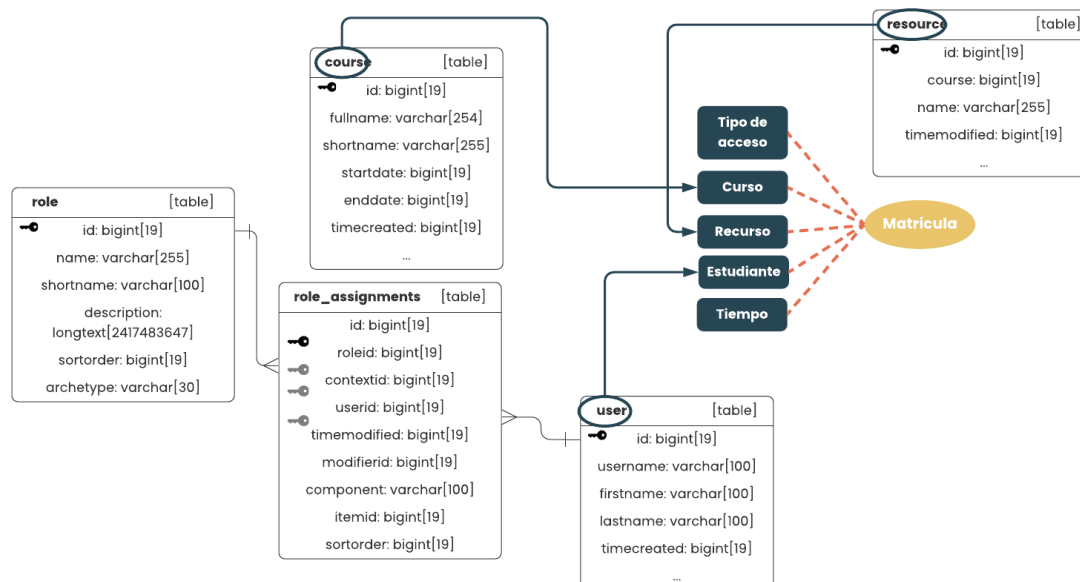
P_SA_1. Se desea conocer la *cantidad de estudiantes* que han enviado a tiempo, tarde o que no han enviado *actividades de tipo assign* en cada una de las diferentes *semanas* planificadas a lo



largo de todo el **curso**. Además, permitir la posibilidad de enviar un email a los estudiantes que han enviado, no han enviado o han enviado tarde una actividad de tipo **assign**.



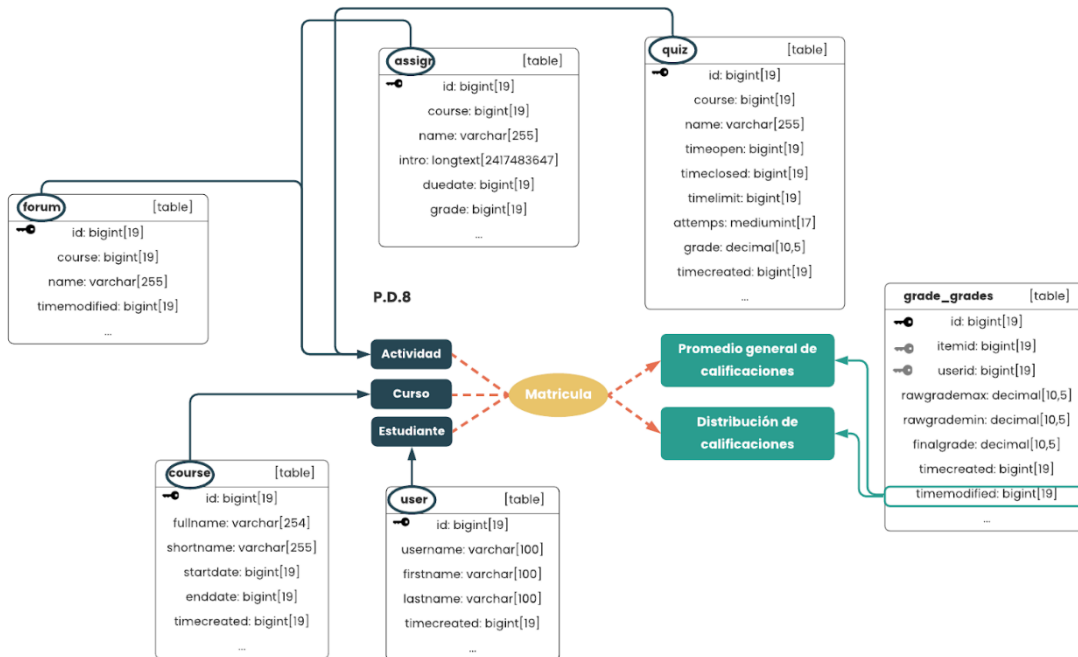
P_SA_2. Se desea conocer la **cantidad de estudiantes** que han accedido y no han accedido a los **recursos y actividades** (que el docente seleccione) durante las diferentes **semanas** de duración del **curso**. Además, permitir la posibilidad de enviar un email a los estudiantes que han accedido o no a dichos recursos y actividades.





Seguimiento (para evaluaciones)

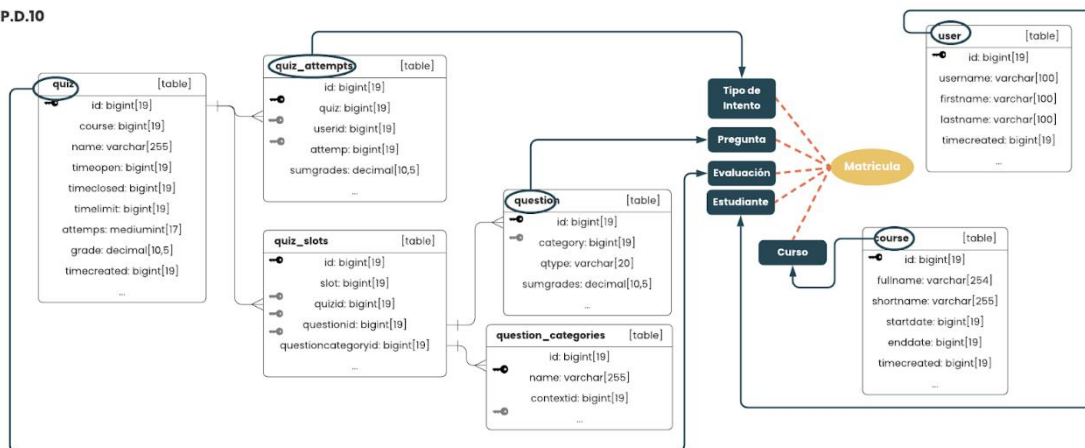
P_SE_1. Se desea conocer el *promedio general de calificaciones* de cada una de las *actividades* evaluables del *curso*. Al seleccionar una actividad específica mostrar la *distribución de calificaciones* con el promedio, la nota máxima y la nota mínima para dicha actividad.



P_SE_2. Se desea conocer la *cantidad de total de intentos, intentos parcialmente correctos, intentos incorrectos, intentos correctos, preguntas en blanco y sin calificar* en cada una de las *preguntas* de una *evaluación* de todos los intentos realizados por los *estudiantes*.



P.D.10



6.1.2.3. Nivel de granularidad

Una vez que se han establecido las relaciones con los OLTP, se deben seleccionar los campos que contendrá cada perspectiva, ya que será a través de estos por los que se examinarán y filtrarán los indicadores. Para ello, basándose en las correspondencias establecidas en el paso anterior, se debe presentar a los usuarios los datos de análisis disponibles para cada perspectiva. Es muy importante conocer en detalle qué significa cada campo y/o valor de los datos encontrados en los OLTP.

De acuerdo a las correspondencias establecidas, se analizaron los campos residentes en cada tabla a la que se hacía referencia analizando la base de datos de Moodle para intuir los significados de cada campo.

- Con respecto a la perspectiva “*Estudiante*” y “*Docente*”, los datos necesarios son los siguientes:
 - *id*: identificador único del usuario en Moodle
 - *username*: nombre del usuario en Moodle
 - *firstname*: nombre del usuario
 - *lastname*: apellido del usuario
 - *timecreated*: indica la fecha y hora de creación de un usuario en un curso de Moodle (en segundos)

Para determinar que un usuario es un estudiante se hace uso de la tabla “*role_assignments*”, los datos necesarios son los siguientes:



- *roleid*: corresponde al identificador del rol en la tabla “**role**” que indica si el usuario es docente o estudiante
- *userid*: corresponde al identificador del usuario registrado en un curso Moodle

Para obtener el nombre del rol del usuario en Moodle se hace uso de la tabla “role”, los datos necesarios son los siguientes:

- *id*: identificador de cada uno de los roles disponibles en Moodle
 - *name*: nombre del rol
- Con respecto a la perspectiva “**Curso**”, los datos necesarios son los siguientes:
 - *id*: identificador único de cada uno de los cursos disponibles en Moodle
 - *fullname*: nombre completo del curso
 - *shortname*: nombre corto o abreviado de un curso
 - *startdate*: fecha de inicio del curso (en segundos)
 - *enddate*: fecha prevista para el final del curso (en segundos)
 - *timecreated*: fecha en la que ha sido creado el curso (en segundos)
 - Con respecto a la perspectiva “**Tipo Actividad**”, los datos necesarios son los siguientes:

Para obtener los diferentes tipos de actividades de un curso se hace uso de la tabla “**course_modules**”, los datos necesarios son los siguientes:

- *course*: identificador del curso al que pertenece una actividad
- *module*: identificador del módulo (tipo de actividad) disponible en un curso
- *instance*: identificador de la instancia del tipo de módulo específico (ejem: si el tipo de módulo es assign, la instancia es el id del registro correspondiente en la tabla mdl_assign)
- *section*: identificador de la sección del curso donde se encuentra disponible el módulo (tipo de actividad)

Con el identificador module de la table “**course_modules**”, se hace referencia a la tabla “**modules**”, los datos necesarios de esta última tabla son los siguientes:



- *id*: identificador del tipo de módulo
- *name*: nombre del módulo (assign, book, chat, choice, data, feedback, folder, forum, etc.)
- Con respecto a la perspectiva “*Evaluación*”, los datos necesarios son los siguientes:
 - *id*: identificador único de una evaluación
 - *course*: identificador del curso donde fue creada la evaluación
 - *name*: nombre de la evaluación
 - *timeopen*: fecha y hora de apertura de la evaluación (en segundos)
 - *timeclosed*: fecha y hora de cierre de la evaluación (en segundos)
 - *timelimit*: tiempo límite para los intentos de la evaluación (en segundos)
 - *attempts*: el máximo número de intentos permitidos a los estudiantes
 - *grade*: el total al que se escala la calificación general de la prueba.

Además de la información anterior sobre una evaluación es necesaria la información de la tabla “**quiz_attempts**”, los datos necesarios son los siguientes:

- *quiz*: identificador de la evaluación al que corresponde el intento
- *userid*: identificador del usuario que realizó el intento en la evaluación
- *attempt*: numera secuencialmente los intentos de este estudiante en este cuestionario.
- *sumgrades*: puntuación total para este intento.

Para conocer los detalles respecto a las *preguntas* contestadas en una evaluación es necesario hacer uso de la tabla “**quiz_slots**” que guarda la información sobre las preguntas de una evaluación:

- *id*: identificador de la pregunta
- *slot*: orden en el que aparece la pregunta
- *quizid*: identificador de la evaluación
- *questionid*: identificador de la pregunta
- *page*: número de página en la que aparece la pregunta
- *maxmark*: los punto que aporta esta pregunta al quiz.sumgrades



- Con respecto a la perspectiva “**Tipo actividad (assign)**”, los datos necesarios son los siguientes:
 - *id*: identificador de page
 - *course*: curso al que pertenece la página
 - *name*: nombre de la instancia de la asignación. Se muestra en la parte superior de cada página.
 - *duedate*: fecha límite para la asignación. Mostrado a los estudiantes.
 - *grade*: calificación máxima para esta tarea.
 - *timemodified*: Hora a la que se modificó por última vez la configuración de esta instancia de módulo de asignación.
 - Con respecto a la perspectiva “**Tipo actividad (folder)**”, los datos necesarios son los siguientes:
 - *id*: identificador de la actividad de tipo carpeta
 - *course*: curso al que pertenece la carpeta
 - *name*: nombre de la instancia de tipo carpeta
 - *timemodified*: fecha y hora de creación (en segundos)
 - Con respecto a la perspectiva “**Tipo actividad (page)**”, los datos necesarios son los siguientes:
 - *id*: identificador de page
 - *course*: curso al que pertenece la página
 - *name*: nombre de la página
 - *timemodified*: fecha y hora de creación (en segundos)
 - Con respecto a la perspectiva “**Tipo actividad (resource)**”, los datos necesarios son los siguientes:
 - *id*: identificador del recurso
 - *course*: curso al que pertenece el recurso
 - *name*: nombre de la instancia de tipo recurso
 - *timemodified*: fecha y hora de creación (en segundos)
 - Con respecto a la perspectiva “**Tipo actividad (url)**”, los datos necesarios son los siguientes:
-



- *id*: identificador de la url
 - *course*: curso al que pertenece la url
 - *name*: nombre de la url
 - *timemodified*: fecha y hora de creación (en segundos)
- Con respecto a la perspectiva “**Actividad**”, se hace uso de la tabla “**grade_items**” almacena todos los ítems evaluables del curso. Los datos necesarios son los siguientes:
 - *courseid*: identificador del curso al que pertenece un ítem evaluable
 - *itemname*: nombre del ítem (introducido por el módulo)
 - *itemtype*: tipo del ítem: 'mod', 'blocks', 'import', 'calculated' etc.
 - *itemmodule*: 'forum', 'quiz', 'csv', etc
 - *iteminstance*: identificador del módulo del ítem (si el tipo de módulo es assign, la instancia es el id del registro correspondiente en la tabla mdl_assign)
 - *grademax*: calificación máxima permitida
 - *grademin*: calificación mínima permitida
 - *timecreated*: fecha y hora de creación del ítem
 - *timemodified*: la última vez que se modificó este ítem
- Con respecto a la perspectiva “**Tipo actividad (forum)**”, los datos necesarios son los siguientes:
 - *id*: identificador del foro
 - *course*: curso al que pertenece el foro
 - *name*: nombre del foro
 - *timemodified*: fecha y hora de creación (en segundos)

6.2. Identificación de Patrones de Comportamiento

Para identificar los patrones de comportamiento que adoptaron los estudiantes se aplicó Minería de Procesos a los datos del registro de eventos del curso en estudio. Moodle registra cada interacción de un usuario en la plataforma en una tabla de base de datos. En esta tabla se puede conocer la acción que realizó el usuario (ver, crear, editar, eliminar, etc), el contenido sobre el que se realizó la interacción, el identificador del usuario que realizó la interacción, la fecha y hora en la que se realizó la interacción, entre otra información menos relevante para nuestro estudio. Además, mediante uniones con otras tablas de la base de datos se puede conocer detalles de la interacción,



como por ejemplo nombre del contenido del curso con el que el usuario interactuó, la sección del curso en la que está el contenido, las configuraciones del contenido con el que se realizó la interacción, los detalles del usuario que realizó la interacción entre otras. Para el caso de este estudio, del conjunto total de atributos de la tabla de interacciones se seleccionaron un conjunto reducido de seis variables (Tabla 7).

Tabla 7*Atributos analizados de la Tabla de Registro de Eventos de Moodle*

component	action	target	object table	object id	user id	course id	time created
core	viewed	course	[null]	[null]	2	2	1598481563
mod_folder	viewed	course_module	folder	3	2	2	1598481524
mod_forum	created	discussion	forum	233	2	2	1598463186
...

Con los atributos seleccionados se generó un nuevo registro de eventos con una mejor descripción de la acción realizada por el estudiante (Tabla 8).

Tabla 8*Log de Eventos con información detallada de la interacción*

Atributo	Descripción
userid	Identificador del usuario.
activity	Nombre de la actividad con la que interactuó
interaction	Acción que se realizó sobre la actividad
section	Sección del curso de la actividad
timecreated	Fecha y hora registrada de la acción
session	Número de la sesión del usuario en la que realizó la interacción.

De todas las interacciones registradas por Moodle, en total se contabilizaron 24 eventos en el curso analizado (Tabla 9).

**Tabla 9***Acciones Consideradas Relevantes*

Acción
assign – created – comment
assign – submitted – submission
assign – updated – submission_status
assign – updated – submission
assign – viewed – assign
assign – viewed – feedback
folder – downloaded – all_files
folder – viewed – folder
forum – created – discussion
forum – created – discussion_subscription
forum – created – post
forum – created – subscription
forum – updated – post
forum – viewed – discussion
forum – viewed – forum
page – viewed – page
quiz – reviewed – attempt
quiz – started – attempt
quiz – submitted – attempt
quiz – viewed – attempt
quiz – viewed – attempt_summary
quiz – viewed – quiz
resource – viewed – resource
url – viewed – url

Finalmente, toda la información antes descrita se almacenó en un archivo separado por comas (csv), mismo que se utilizó como entrada para el software Disco Fluxicon, en el que se graficaron mapas de procesos con el comportamiento de los estudiantes en cada semana y en cada sección del curso.

Como fue estructurado el curso, debía haber una correspondencia entre las actividades realizadas de los estudiantes en cada semana con cada sección del curso, como se muestra en la Tabla 10. Por lo tanto, el análisis se realizó en base al tiempo y en a las interacciones con el objetivo de



encontrar si la correspondencia antes mencionada se cumplió o si hubo variantes al comportamiento esperado por parte de los estudiantes.

Tabla 10

Planificación en Semanas del Curso Planificación y Evaluación de Clases Virtuales

Semana	Duración	Sección del Curso
0	antes del 25/05/2020	Sección 0 (Bienvenida)
1	desde el 25/05/2020 hasta el 31/05/2020	Sección 1 (Semana 1 – Introducción)
2	desde el 01/06/2020 hasta el 07/06/2020	Sección 2 (Semana 2 – El Tutor y la Comunicación)
3	desde el 08/06/2020 hasta el 14/06/2020	Sección 3 (Semana 3 – Los Materiales Educativos)
4	desde el 15/06/2020 hasta el 21/06/2020	Sección 4 (Semana 4 – La Evaluación de los Aprendizajes)
5	después del 21/06/2020	Sección 5 (Evaluación del Curso)

En primer lugar, se analizaron las interacciones que sucedieron en la Semana 0 (antes del 25 de mayo), en la que los estudiantes debían haber trabajado únicamente con los contenidos de la Sección 0 (Bienvenida del Curso), la cual contiene dos foros, una evaluación, dos recursos tipo url y un recurso tipo pdf. Para este caso se generaron dos mapas de procesos: el primero con las interacciones realizadas por los estudiantes antes del 25 de mayo (en base al tiempo, Figura 31 a) y el segundo con las interacciones de los estudiantes con los contenidos de la sección de bienvenida del curso (en base a las interacciones, Figura 31 b).

Como se puede observar, el comportamiento de los estudiantes en la semana 0 que se evidencia ambos mapas de procesos (Figura 31) muestra que interactuaron con los contenidos de la sección de Bienvenida del curso, como se esperaba. La principal diferencia es que, en el mapa de procesos de las interacciones (Figura 31 b), se observa, en base a las frecuencias, que casi todos los estudiantes realizaron la evaluación de la sección, lo cual no se aprecia en el mapa de procesos del tiempo (Figura 31 a). Esto implica que, si bien todos los estudiantes realizaron la evaluación de la sección, la mayoría no lo hizo en la primera semana como se esperaba.

La siguiente semana de análisis fue la Semana 1 (25 de mayo - 31 de mayo), en la que los estudiantes debían haber trabajado únicamente con los contenidos de la Sección 1 (Introducción), la cual contiene tres foros, una tarea, un recurso tipo word, cuatro recursos tipo url y ocho recursos tipo pdf. Al igual que la semana 0, se generaron dos mapas de procesos: el primero con las interacciones realizadas por los estudiantes entre el 25 y el 31 de mayo (en base al tiempo, Figura 32 a) y el segundo



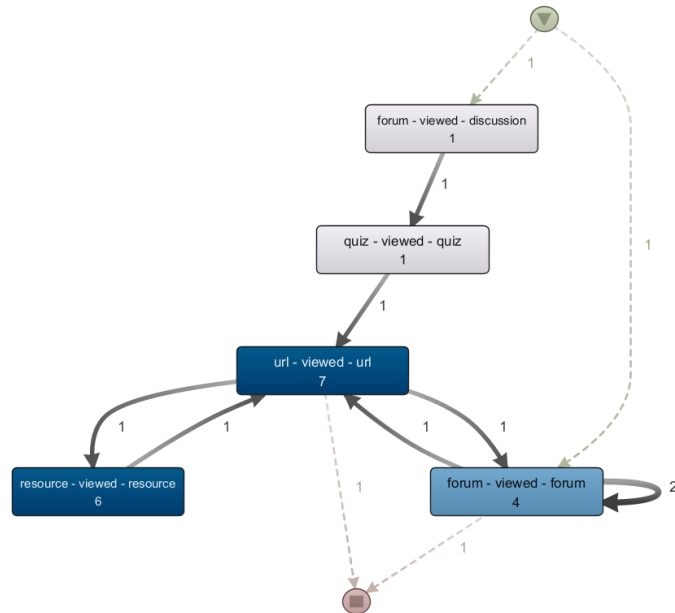
con las interacciones de los estudiantes con los contenidos de la sección 1 (en base a las interacciones, Figura 32 b).

En los contenidos de la semana 1 del curso existe una gran cantidad de recursos y foros, por lo que se esperaría que las interacciones de los estudiantes estén relacionadas a las actividades de los tipos forum, resources y assign. Tales comportamientos se evidencian en los mapas de procesos generados de esta sección (Figura 32), en la que se puede observar que principalmente los estudiantes acceden a los foros (forum – view – forum, forum – view – discussion) y ven los recursos (resource – viewed – resource). También existen interacciones relacionadas a tareas (assign – uploaded – submission, assign – submitted – submission) en ambos mapas de procesos, pero en menor cantidad.

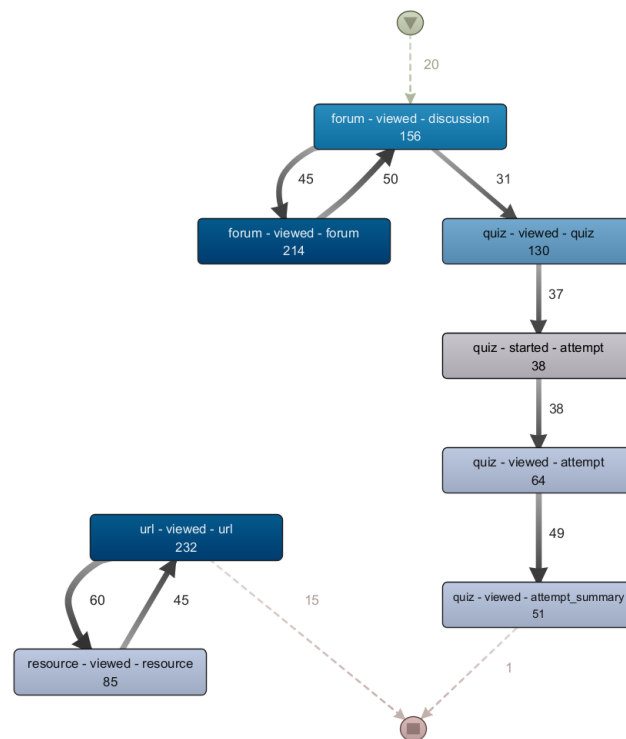


Figura 31

Mapas de Procesos de la Semana 0



a) Mapa de Procesos por Semana

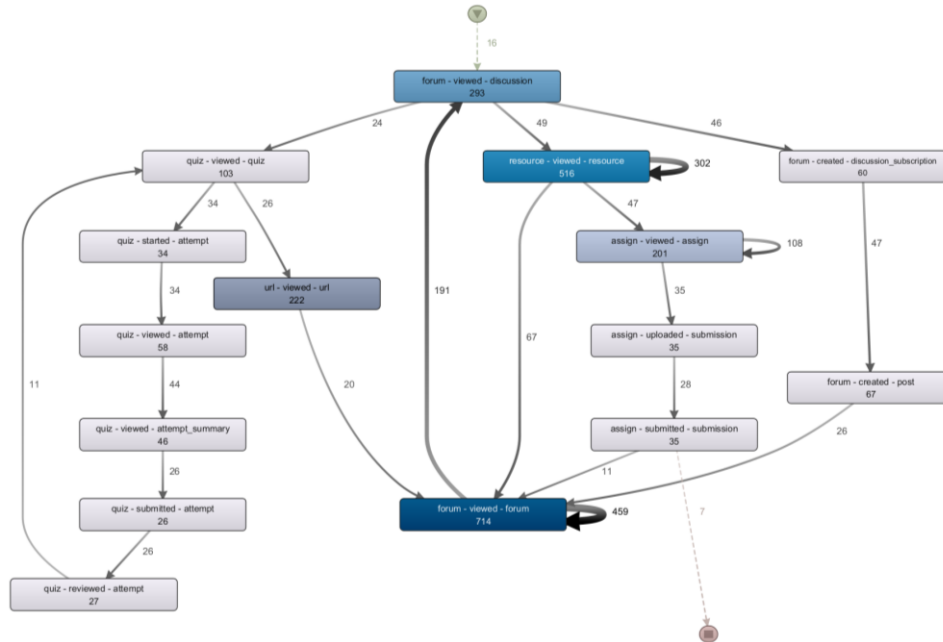


b) Mapa de Procesos por Sección

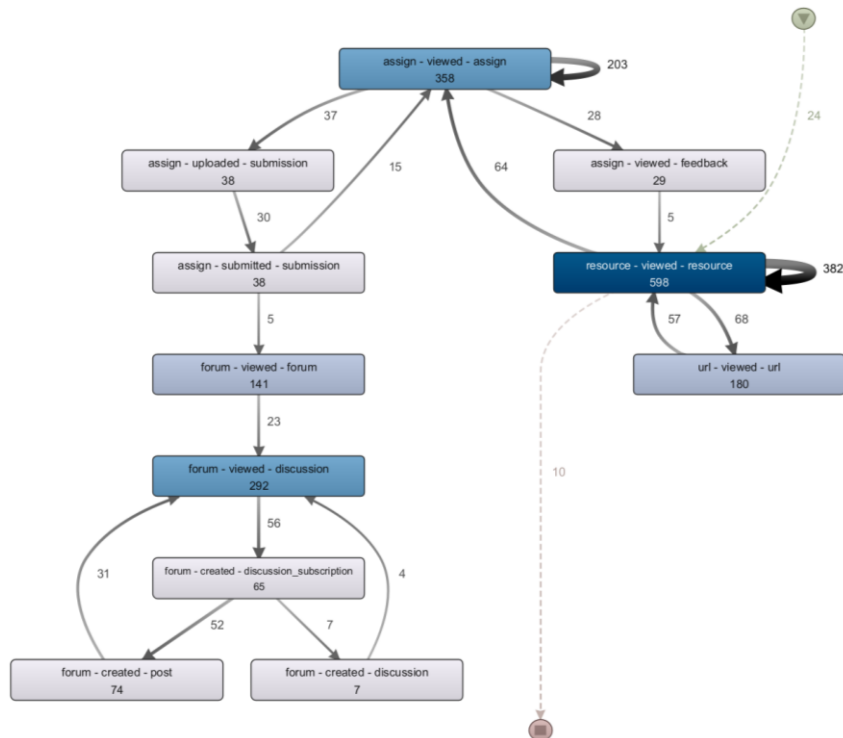


Figura 32

Mapas de Procesos de la Semana 1



a) Mapa de Procesos por Semana



b) Mapa de Procesos por Sección

La diferencia encontrada es que, en el mapa de procesos del tiempo (Figura 32 a), se observa que existen interacciones relacionadas a una evaluación (quiz – start – attempt, quiz – view – attempt, quiz – submitted – attempt), a pesar de que en la sección no existe una evaluación. Esto implica que varios de los estudiantes realizaron la evaluación de la sección de bienvenida del curso en esta semana, lo cual coincide con el comportamiento observado en el mapa de procesos del tiempo de la semana 0 (Figura 31 a), en donde se observa que los estudiantes solo acceden a la evaluación, mas no la resuelven.

La siguiente semana de análisis fue la Semana 2 (01 de junio - 07 de junio), en la que los estudiantes debían haber trabajado únicamente con los contenidos de la Sección 2 (El Tutor y la Comunicación), la cual contiene dos foros, una tarea, un recurso tipo word, dos recursos tipo excel, un recurso tipo url, siete recursos tipo pdf, un recurso tipo página y dos carpetas. De la misma forma que las semanas anteriores, se generaron dos mapas de procesos: el primero con las interacciones realizadas por los estudiantes entre el 01 y el 07 de junio (en base al tiempo, Figura 33 a) y el segundo con las interacciones de los estudiantes con los contenidos de la sección 2 (en base a las interacciones, Figura 33 b).

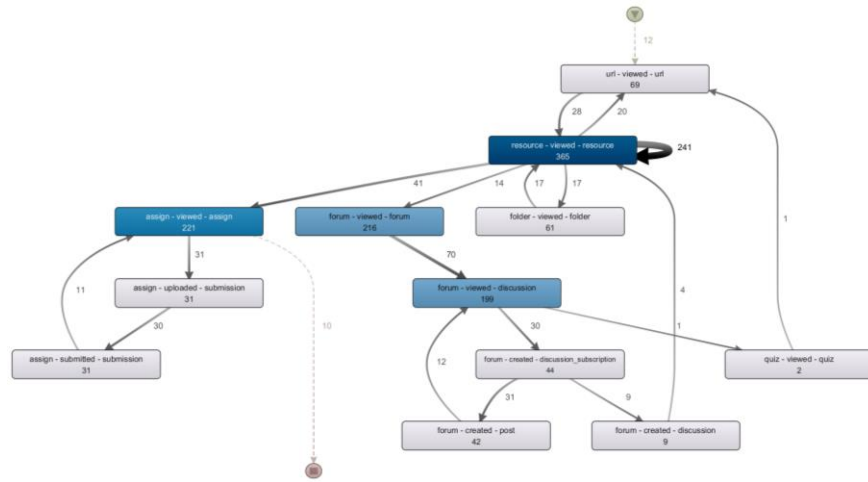
A diferencia de las anteriores, en esta semana ya se puede evidenciar que el comportamiento del estudiante empieza a ser el esperado. Los dos mapas de procesos para esta semana (Figura 33) muestran que los estudiantes realizan con gran frecuencia interacciones relacionadas a los recursos, foros y carpetas de la sección (url – viewed – url, resource – viewed – resource, forum – viewed – forum, forum – created – post, etc). Sin embargo, en el mapa de proceso del tiempo (Figura 33 a) existen interacciones relacionadas con una evaluación, a pesar de que ni en esta sección ni en la anterior existe una evaluación. Este comportamiento indica que, inclusive hasta la tercera semana existieron estudiantes que resolvieron la evaluación de la primera semana.

La siguiente semana de análisis fue la Semana 3 (08 de junio - 14 de junio), en la que los estudiantes debían haber trabajado únicamente con los contenidos de la Sección 3 (Los Materiales Educativos), la cual contiene dos foros, cuatro recursos tipo url, cinco recursos tipo pdf y dos recursos tipo página. De la misma forma que las semanas anteriores, se generaron dos mapas de procesos: el primero con las interacciones realizadas por los estudiantes entre el 08 y el 14 de junio (en base al tiempo, Figura 34 a) y el segundo con las interacciones de los estudiantes con los contenidos de la sección 2 (en base a las interacciones, Figura 34 b).

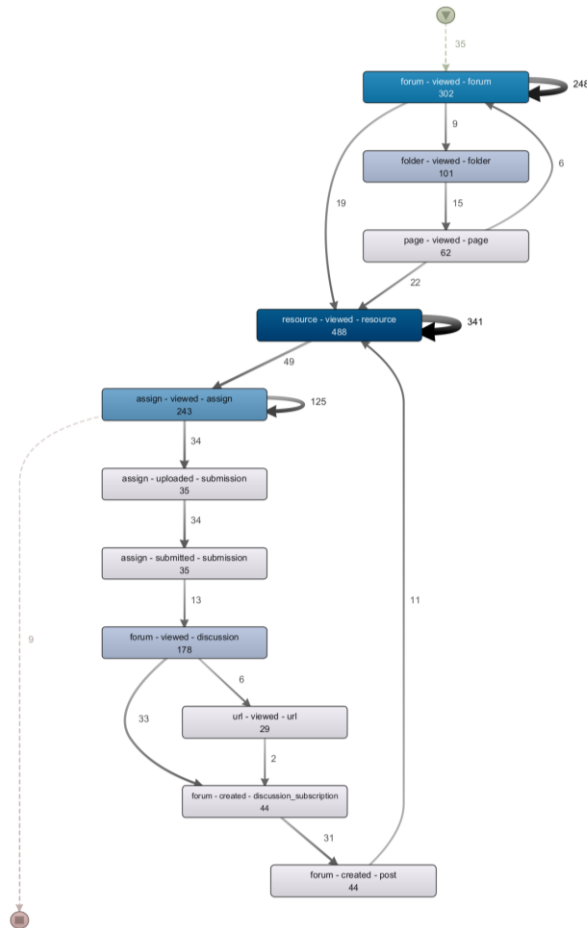


Figura 33

Mapas de Procesos de la Semana 2



a) Mapa de Procesos por Semana

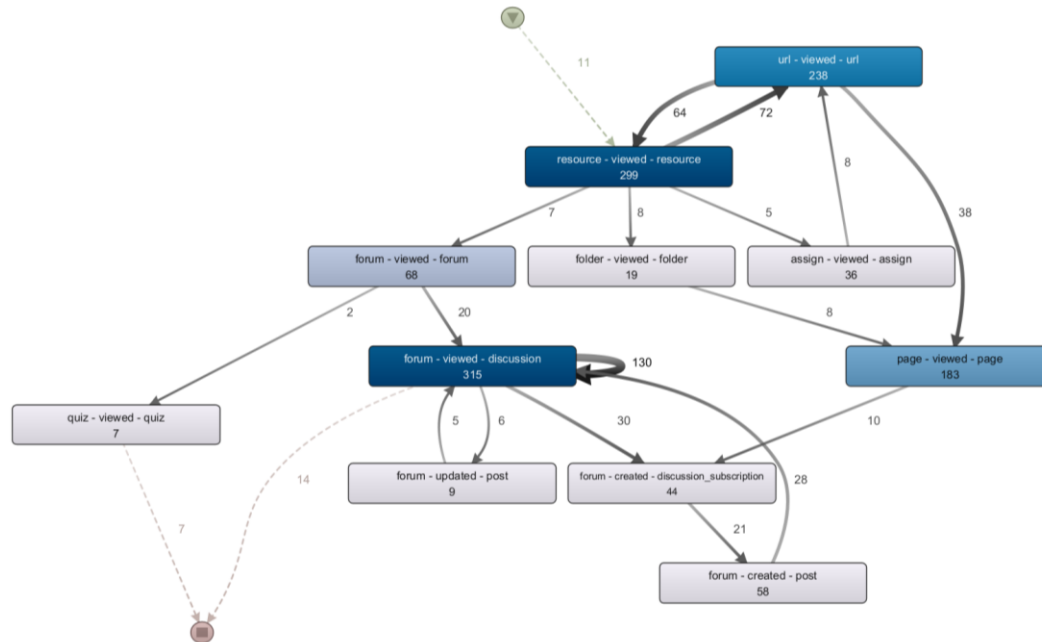


b) Mapa de Procesos por Sección

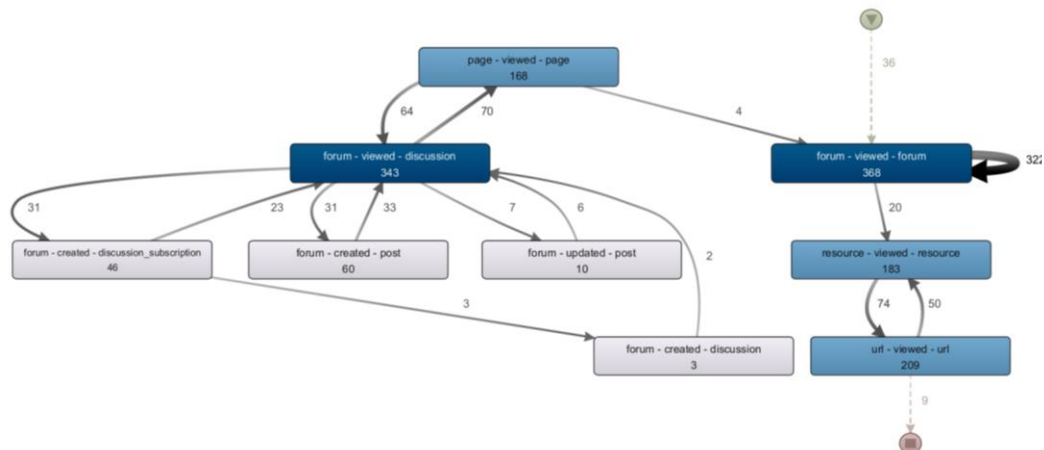


Figura 34

Mapas de Procesos de la Semana 3



a) Mapa de Procesos por Semana



b) Mapa de Procesos por Sección

Analizando los mapas de procesos generados (Figura 34), vemos que en esta semana los estudiantes tuvieron un comportamiento muy similar a la anterior. Hay muchas interacciones con los foros y los recursos del curso (forum – viewed – forum, resource – viewed – resource, etc). Inclusive se observa que aún existen estudiantes que interactúan con la evaluación de la primera semana (Figura 34 a), tal como en la semana anterior.



La siguiente semana de análisis fue la Semana 4 (15 de junio - 21 de junio), en la que los estudiantes debían haber trabajado únicamente con los contenidos de la Sección 4 (La Evaluación y los Aprendizajes), la cual contiene dos foros, siete recursos tipo url, cinco recursos tipo pdf, dos carpetas y una evaluación. De la misma forma que las semanas anteriores, se generaron dos mapas de procesos: el primero con las interacciones realizadas por los estudiantes entre el 15 y el 21 de junio (en base al tiempo, Figura 35 a) y el segundo con las interacciones de los estudiantes con los contenidos de la sección 3 (en base a las interacciones, Figura 35 b).

Al igual que las semanas anteriores, los estudiantes mantienen una alta frecuencia de interacción con los contenidos y con los foros de la sección. Algo para resaltar en esta semana es la presencia de una evaluación (Figura 35), la cual se ve reflejada en las interacciones que han hecho los estudiantes (quiz – viewed – quiz, quiz – started – attempt, quiz – submitted – attempt).

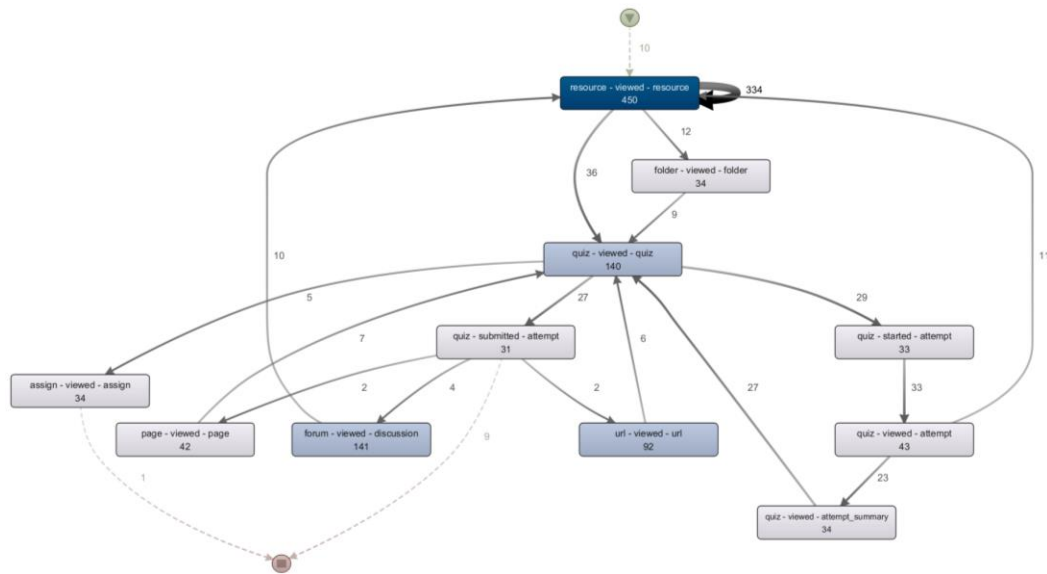
La última semana de análisis fue la Semana 5 (después del 21 de junio), en la que los estudiantes debían haber trabajado únicamente con los contenidos de la Sección 5 (Evaluación del curso), la cual contiene un recurso tipo url y un recurso tipo página. De la misma forma que las semanas anteriores, se generaron dos mapas de procesos: el primero con las interacciones realizadas por los estudiantes después del 21 de junio (en base al tiempo, Figura 36 a) y el segundo con las interacciones de los estudiantes con los contenidos de la sección 3 (en base a las interacciones, Figura 36 b).

En esta semana se dió un caso particular, ya que su sección contenía únicamente dos recursos, tal como se ve en el mapa de procesos de la Figura 36 b. Sin embargo, las interacciones para esta semana se contabilizaron desde que terminó la cuarta semana, es decir a partir del 21 de junio, por lo que esta semana registró toda la actividad de los estudiantes una vez terminado el curso. Como se puede observar en el mapa de procesos del tiempo (Figura 36 a), en esta semana los estudiantes realizaron interacciones relacionadas a resolución de evaluaciones (quiz – started – attempt, quiz – submitted – attempt, etc), envíos de tareas (assign – uploaded – submission, assign – submitted – submission, etc), publicaciones en foros (forum – viewed – discussion, forum – created post, etc), entre otras. Esto indica que los estudiantes utilizaron la última semana del curso para cumplir con todas las actividades pendientes y enviarlas para la calificación del docente.

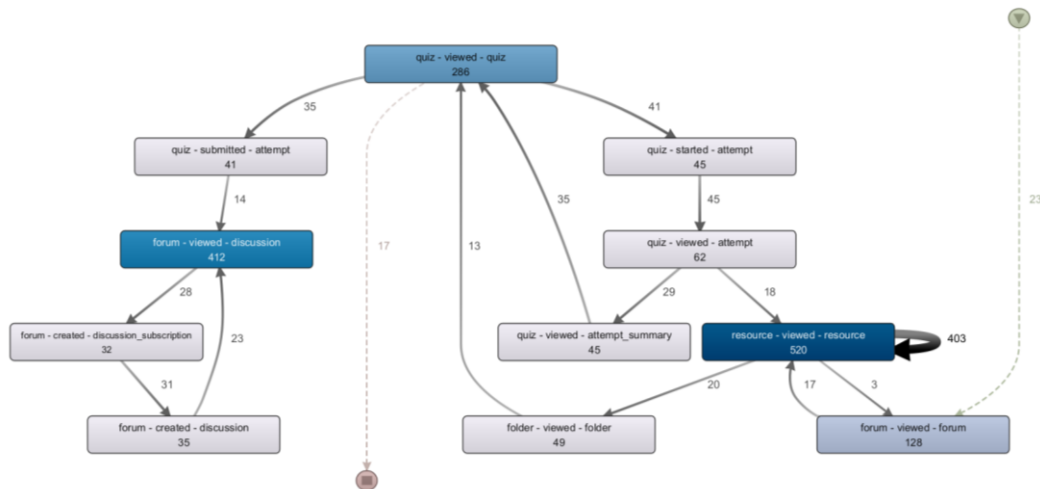


Figura 35

Mapas de Procesos de la Semana 4



a) Mapa de Procesos por Semana

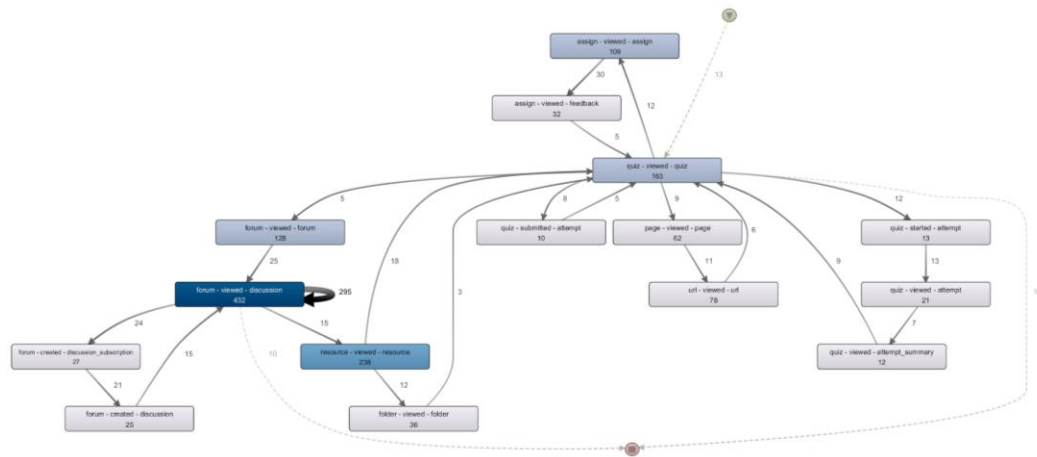


b) Mapa de Procesos por Sección

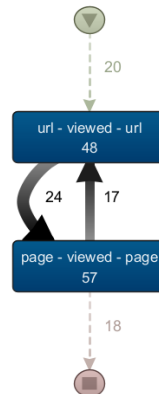


Figura 36

Mapas de Procesos de la Semana 5



a) Mapa de Procesos por Semana



b) Mapa de Procesos por Sección

Luego del análisis de todas las semanas se pudo evidenciar que los estudiantes empiezan el curso sin tener un comportamiento esperado, luego normalizan sus actividades en el transcurso de las semanas hasta antes del final del curso, en donde el comportamiento vuelve a diferir del esperado. Así, se evidenció un comportamiento no esperado de los estudiantes en las semanas 0, 1 y 5 del curso. Mientras que, en las semanas 2, 3 y 4 los estudiantes se comportaron como se esperaba que lo hicieran. Algunos detalles que vale destacar es que al inicio del curso, algunos estudiantes no realizaron una evaluación inicial del curso hasta la tercera semana del curso. Además, la semana final del curso sirvió para que la mayoría de los estudiantes cumplieran con las actividades que tenían pendientes. Además, varios de ellos acceden a recursos a destiempo, es decir, a pesar de que el curso presenta



una estructura bien definida, dichos estudiantes no cumplen con el plan académico propuesto por el docente. Muchos de estos comportamientos pueden traducirse como falta de compromiso, conllevar a problemas de aprendizaje e incluso riesgo de deserción.

(Figura 49).

6.3. Detección de Variables de Predicción de Abandono

Hasta la fecha, las contribuciones de investigación en el campo de la analítica del aprendizaje han tenido como objetivo general ampliar la comprensión del proceso de aprendizaje (Pardo, y otros, 2016). La tasa de deserción excesiva en cursos virtuales ha animado a los investigadores a pensar en el uso de métodos de análisis de aprendizaje (Xing & Du, 2018) para la predicción temprana de los estudiantes que están en riesgo de deserción.

Un aspecto central de este estudio ha sido el desarrollo de un modelo predictivo de comportamiento, rendimiento y riesgo de deserción de los estudiantes. Los modelos comúnmente ayudan a identificar posibles relaciones entre varios factores conductuales, demográficos y basados en el rendimiento y compromiso del estudiante en conjuntos de datos (Romero, Ventura, & Garcia, 2008). Si bien estos modelos predictivos se han utilizado para desarrollar indicadores tempranos, se ha prestado poca atención a la investigación de cómo dicha información se puede implementar mejor para promover la reflexión y la acción entre docentes y estudiantes.

Para lograr construir un modelo predictivo robusto, se ha tomado como punto de partida al modelo proporcionado por Moodle. Este modelo brinda un conjunto de indicadores que se basan en la participación de los estudiantes en el curso. Los indicadores pueden ser de dos tipos: Profundidad Cognitiva (*CD – Cognitive Depth en inglés*) y amplitud social (*SB – Social Breadth en inglés*).

6.3.1. Profundidad Cognitiva

La profundidad cognitiva se define como “La medida en que los participantes en cualquier configuración particular de una comunidad de investigación son capaces de construir significado a través de una comunicación sostenida” (Garrison, Anderson, & Archer, 2000). La profundidad cognitiva generalmente se determina mediante análisis de contenido. Para definir indicadores de profundidad cognitiva, se toma como base el tipo de actividad que se le ofrece al alumno y en qué medida el alumno demuestra un compromiso cognitivo en esa actividad (Moodle, 2020). El nivel de profundidad varía de 0 a 5, donde 0 indica que el alumno ni siquiera ha visto la actividad. Los niveles de profundidad cognitiva potencial son (Moodle, 2020):



- El estudiante ha visto los detalles de la actividad.
- El estudiante ha enviado contenido a la actividad.
- El estudiante ha visto los comentarios de un instructor o un compañero para la actividad (incluidos los comentarios automatizados).
- El estudiante ha proporcionado comentarios al instructor o un compañero dentro de la actividad.
- El estudiante ha revisado y/o vuelto a enviar contenido a la actividad.

El modelo comienza asignando un valor potencial máximo de profundidad cognitiva a cada módulo de actividad. Por ejemplo, el módulo de Tareas (**assign**) permite una profundidad cognitiva de 5. En la Tabla 11 se muestra la asignación de niveles para las actividades de Moodle.

Una vez asignados los niveles potenciales, se evalúa a cada alumno inscrito en un curso en función de la proporción de profundidad cognitiva alcanzada. Por ejemplo, si una actividad solo admite hasta el nivel 3 (visualización de comentarios) y el alumno ha alcanzado el nivel 3, el alumno participa al 100 por ciento del nivel posible de profundidad cognitiva y el valor del indicador es 1. Si el alumno no ve la actividad, tienen una profundidad cognitiva de 0 y el valor del indicador es -1. Si el estudiante ha visto la actividad y ha enviado una respuesta, pero no ha visto comentarios, se encuentra en el nivel de profundidad cognitiva 2 (fuera de una profundidad potencial en esta actividad de 3), por lo que el indicador se calcula en .33.

Tabla 11

Asignación de Niveles de Profundidad Cognitiva para las diferentes Actividades de Moodle

Nivel	Ponderación	Actividad
1	0% (no visto) o 100% (visto)	Libro, Carpeta, Paquete IMS, Etiqueta, Página, Archivo, URL
2	0% (no visto) 50% (visto) o 100% (interacción con otro participante)	Opción, Base de Datos, Retroalimentación, Glosario, Encuesta, Wiki
3	0% (no visto) 33% (visto) 67% (enviado) o 100% (comentario enviado y visto)	Opción, LTI, SCORM
4	0% (no visto) 25% (visto) 50% (enviado) 75% (comentarios enviados y vistos) o 100% (comentarios al docente o compañero).	Chat, Foro



	0% (no visto)	
	20% (visto)	
	40% (enviado)	
5	60% (comentarios enviados y vistos)	Tarea, Lección, Quiz, Taller
	80% (comentarios al docente o compañero) o	
	100% (revisado y reenviado)	

6.3.2. *Amplitud social*

La amplitud social es una medida de "presencia social". Se define como “La capacidad de los participantes para identificarse con el grupo o curso de estudio, comunicarse intencionalmente en un entorno de confianza y desarrollar relaciones personales y afectivas progresivamente mediante la proyección de sus personalidades individuales” (Garrison, 2009). Este modelo implementa la presencia social como "amplitud social" al examinar la amplitud de oportunidades que tiene el participante (estudiante) para comunicarse con los demás (Moodle, 2020). El nivel de amplitud varía de 0 a 2, donde 0 indica que el alumno no ha interactuado con nadie. Los niveles de amplitud social potencial son (Moodle, 2020):

- El estudiante no ha interactuado con ningún otro participante en esta actividad.
- El estudiante ha interactuado con al menos otro participante (por ejemplo, ha enviado una tarea o ha intentado una prueba proporcionando comentarios).

Este modelo comienza asignando un valor potencial máximo de amplitud social a cada módulo de actividad. Por ejemplo, el módulo Asignación permite una amplitud social de 2. En la Tabla 12 se muestra la asignación de niveles para las actividades de Moodle. Una vez asignados los niveles potenciales, se evalúa a cada alumno inscrito en un curso en función de la proporción de profundidad potencial alcanzada. Por ejemplo, si una actividad solo admite hasta el nivel 3 y el estudiante ha alcanzado el nivel 3, el estudiante participa al 100 por ciento del nivel posible de amplitud social.

Moodle en su versión 3.9 solo proporciona indicadores hasta el nivel de amplitud social 2. Todos los "recursos" tienen una amplitud social máxima de 1. Las actividades varían en su amplitud social admitida. Esta función está destinada a ampliarse en futuras versiones de Moodle.

Tabla 12

Asignación de Niveles de Amplitud Social para las diferentes Actividades de Moodle

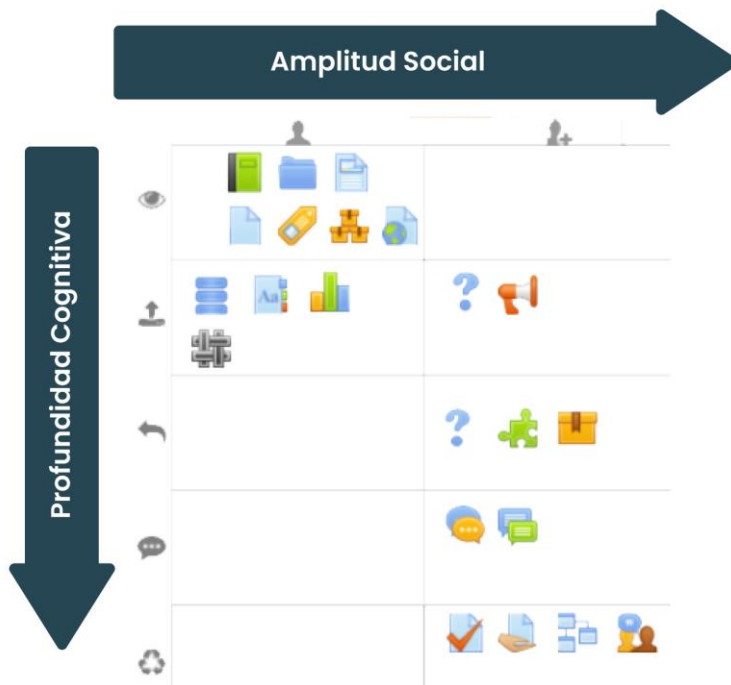
Nivel	Ponderación	Actividad
1	0% (no visto) o 100% (visto)	Libro, Base de Datos, Carpeta, Glosario, Paquete IMS, Etiqueta, Página, Archivo, Encuesta, URL, Wiki
2	0% (no visto) 50% (visto) o 100% (interacción con otro participante)	Tarea, Chat, Elección, Retroalimentación, Foro, Lección, LTI, Quiz, SCORM, Taller

6.3.3. Niveles de indicadores para Módulos de Actividad

El potencial de participación a través de la profundidad cognitiva y la amplitud social constituye el diseño instruccional, que es un elemento clave de la presencia docente. La Figura 37 muestra la profundidad cognitiva y la amplitud social de las actividades y recursos de Moodle.

Figura 37

Matriz de Profundidad Cognitiva y Amplitud Social de Actividades y Recursos de Moodle



Nota. Adaptada de (Moodle, 2020).



Al clasificar cada actividad o recurso según la profundidad cognitiva y amplitud social, se puede anticipar qué nivel de participación se apoya (y posiblemente se espera) del alumno, incluso sin un historial de las acciones de muchos alumnos en una actividad. Hay que tener en cuenta que los niveles superiores a lo largo de cada eje incluyen todos los niveles inferiores, es decir, una actividad que involucra a un estudiante y todos los compañeros (amplitud social 3) incluye automáticamente los niveles 1 (solo estudiante) y 2 (estudiante +1). En este estudio, el nivel específico se ha determinado analizando la configuración de los parámetros para la actividad.

6.3.4. Variables de Predicción del Modelo

Las variables utilizadas como predictores para construir el modelo se calcularon en base a tres enfoques: el primero fue el tiempo invertido por el estudiante en el curso. El segundo enfoque consideró la interacción de los estudiantes con los diversos recursos y actividades de diseño de aprendizaje disponibles en el curso. Y el tercer enfoque incluyó las interacciones del estudiante con sus compañeros del curso y con el docente. Estos tres enfoques permitieron generar indicadores en base al tiempo invertido (sesiones de estudio), la profundidad cognitiva (interacciones con recursos) y la amplitud social (interacciones entre estudiantes) de las actividades y recursos de Moodle.

6.3.4.1. Indicadores de Tiempo Invertido

El enfoque del Tiempo Invertido permitió generar tres indicadores: el *Número de Sesiones* de estudio, el *Tiempo Invertido* en el curso y los *Días Activos* del estudiante en la plataforma.

Para conocer el número de sesiones se hace uso del término *sesión de estudio*. La definición de sesión implica que cada sesión de estudio tiene un comienzo y un final claros. Sin embargo, en Moodle las sesiones no tienen un formato bien definido en los registros de eventos de la plataforma. Estos registros simplemente almacenan una interacción del estudiante con algún contenido del curso. Cada registro especifica la fecha y hora en la que sucedió la interacción en un formato de cantidad de segundos transcurridos desde la medianoche en el Tiempo Universal Coordinado (*UTC – Coordinated Universal Time en inglés*) del 1 de enero de 1970. Al no contar con una forma de definir claramente una sesión por parte de Moodle, dependió de los investigadores determinar cuándo comienza una sesión y cuándo termina. Para esto, se han tenido en cuenta investigaciones previas que han utilizado diferentes enfoques para elegir el tiempo entre sesiones, tales como el manejo de datos (por ejemplo, análisis de distribución de tiempos entre sesiones) y el aprendizaje basado en el diseño; (Einhardt, Aires Tavares, & Cechinel, 2016); (Romero, Ventura, & Garcia, 2008); (Luna, Castro, & Romero, 2017); (Moreno Marcos, Muñoz Merino, Alario Hoyos, Estévez Ayres, & Delgado Kloos,



2018); (Ginda, Suri, Bueckle, & Börner, 2016). En el estudio actual, se combinó un enfoque basado en datos teniendo en cuenta el diseño de aprendizaje de Moodle.

Para determinar la definición de sesión, se han considerado períodos de inactividad entre las interacciones de los estudiantes dentro de un curso de Moodle. Esta idea se ha basado en la investigación de (G. de Barba, y otros, 2020) en la cual se define una sesión basada en la duración del período entre actividades y distribuciones de frecuencia usando análisis estadístico. La investigación contempla el hecho de que un alumno puede estar activo, pero sin ninguna interacción con el sistema (es decir, viendo o reflexionando sobre el material didáctico). En dicha investigación se define en 30 minutos como el límite de tiempo de inactividad entre interacciones de un estudiante. Eso significa que cualquier par de interacciones cuya diferencia en el tiempo en el que ocurrió es menor a 30 minutos se consideran parte de la misma sesión. De otra forma, si el tiempo entre dos interacciones es mayor a 30 minutos, se consideran diferentes sesiones. Con esta definición de sesión, se calcularon los siguientes indicadores:

- *Número de sesiones*: como ya se explicó, las sesiones se identificaron en los registros de interacciones de los alumnos en función de los accesos al curso y en base a un umbral de 30 minutos entre interacciones.
- *Tiempo invertido*: representa la cantidad de tiempo en minutos que el estudiante permaneció activo en la plataforma. Esta variable se calcula sumando la duración de cada sesión de estudio que realizó en el curso.
- *Días Activo*: representa la cantidad de días que el estudiante realizó al menos una sesión de estudio desde que inició el curso.

6.3.4.2. Indicadores de Profundidad Cognitiva

Como se indicó en secciones anteriores, la profundidad cognitiva está relacionada con las interacciones del estudiante con los contenidos del curso. De esta manera, la profundidad cognitiva se define a nivel de cada contenido individual del curso, por lo que la cantidad de indicadores que este enfoque permite obtener es igual al número de contenidos que tenga un curso. Es necesario mencionar que entre los diversos tipos de contenidos que permite crear Moodle existe el recurso Etiqueta (*label*), el cual no fue considerado para obtener indicadores de Profundidad Cognitiva. Esta decisión se tomó ya que este tipo de recurso siempre presentaba el mismo valor para cada uno de los estudiantes de un curso. Además, en este estudio se calcularon indicadores de profundidad

cognitiva únicamente para los recursos que permite crear Moodle en una instalación estándar. Los recursos de tipo módulo que pueden ser agregados mediante plugins de terceros no fueron tomados en cuenta.

6.3.4.3. Indicadores de Amplitud Social

Los indicadores de Amplitud Social tuvieron un proceso de generación muy similar a los de Profundidad Cognitiva. De igual manera que los indicadores anteriores, no se tomó en cuenta a el recurso Etiqueta (`label`) ni a los recursos tipo módulos externos a los de una instalación estándar de Moodle.

Los datos utilizados para crear un modelo requieren un procesamiento previo sustancial para cumplir con los requisitos impuestos por los algoritmos de Machine Learning. Teniendo en cuenta la generación de indicadores y variables predictoras, el modelo de datos para la predicción se conforma de una matriz conformada por filas y columnas. La cantidad de filas corresponde a la cantidad de alumnos matriculados en el curso, mientras que, la cantidad de columnas depende de la cantidad de módulos, (recursos y actividades del curso), número de sesiones, tiempo invertido, cantidad de días que el estudiante estuvo activo en la plataforma y el identificador del estudiante. Como un ejemplo de cálculo de indicadores de Profundidad Cognitiva y Amplitud Social, si un curso tiene 90 recursos, la cantidad de indicadores calculados será de 183 para cada estudiante, 90 indicadores relacionados a la profundidad cognitiva, 90 relacionados a la amplitud social, un indicador para el número de sesiones, un indicador para el tiempo invertido y un indicador para la cantidad de días activo en la plataforma. Esto daría como resultado una matriz de 184 columnas, la primera de ellas, el identificador del estudiante.

Los indicadores se calcularon en base a las fechas de inicio y finalización del curso. En los casos de cursos que no tienen una fecha de finalización configurada, los indicadores se generaron con todos los datos disponibles desde la fecha de inicio del curso.

6.3.5. Clustering

El análisis de datos de registro o de seguimiento de bajo nivel recopilados automáticamente mientras los estudiantes interactúan con el curso permiten construir modelos de predicción utilizando algoritmos de Machine Learning (Romero, Ventura, & Garcia, 2008), lo que lleva al entrenamiento de un modelo matemático para hacer una detección temprana de deserción estudiantil basado en su comportamiento de interacción previa durante el curso.



El modelo de predicción propuesto se basa en un aprendizaje no supervisado. Se ha hecho uso de un aprendizaje no supervisado ya que no se tiene un conjunto de datos previo que pueda ser tomado como conjunto de datos de entrenamiento con datos etiquetados sobre la aprobación o deserción de un estudiante.

En el aprendizaje electrónico, la agrupación se puede utilizar para encontrar grupos de estudiantes con características de aprendizaje similares y para promover el aprendizaje colaborativo basado en el grupo, así como para proporcionar un diagnóstico de aprendizaje incremental (G. de Barba, y otros, 2020), para descubrir patrones que reflejan los comportamientos del usuario y para gestión de la colaboración para caracterizar grupos de comportamiento similares en espacios de colaboración (Gaudioso & Talavera Méndez, 2005), para agrupar estudiantes con el fin de brindarles una orientación diferenciada según a sus habilidades y otras características (Hämäläinen, Laine, & Sutinen, 2006).

Se ha hecho uso del algoritmo de K-Means (explicado en la sección de [Machine Learning](#)) para generar un modelo predictivo de agrupación (clustering) en base a la matriz de datos explicada en la sección anterior. El objetivo es agrupar a los estudiantes de un curso específico en diferentes grupos (clusters) dependiendo de las actividades realizadas, recursos revisados, tiempo invertido, sesiones de estudio y días activos en la plataforma.

La matriz de datos generada necesita un proceso previo de normalización. La normalización es el proceso de escalar muestras individuales para tener una norma unitaria. Este proceso es útil para cuantificar la similitud de un conjunto de muestras (Arkadiusz, 2020). En otras palabras, la matriz de datos normalizada contendrá valores entre en el rango de 0a 1 y ningún valor será superior a la unidad.

6.3.5.1. Aplicación de K-Means

Para la aplicación del algoritmo de K-Means en primer lugar se debe seleccionar un número de k que será la cantidad de clústeres que generará el algoritmo. La selección del valor de k puede ser hecha de manera aleatoria si se cuenta con la suficiente información previa sobre los datos. De otra forma, existen métodos que permiten estimar un valor de k óptimo en base a un análisis previo de los datos a clusterizar. Para el caso de este estudio se utilizaron los algoritmos “Elbow Curve” y “Silhouette” con el fin comparar los valores de k óptimos que producen estos algoritmos.

Para la aplicación de tanto de “Elbow Curve” como de “Silhouette” en primer lugar es necesario tener datos de los indicadores de tiempo invertido, profundidad cognitiva y amplitud social de los estudiantes del curso a analizar. Para el estudio en cuestión se ha utilizado una muestra de 40 estudiantes pertenecientes a un curso con 70 recursos de diferentes tipos. El proceso de generación



de indicadores (descrito en la sección anterior) da como resultado una matriz de 40 columnas x 148 filas, como se puede observar en la Tabla 13:

Tabla 13

Indicadores de Tiempo Invertido, Profundidad Cognitiva (CD) y Amplitud Social (SB)

Usuario	Tiempo invertido	Número sesiones	Días activo	Recurso 1 (CD)	Recurso 2 (CD)	...	Recurso1 (SB)	Recurso 2 (SB)	...
4	512.8	41	21	1	0.5	...	0	0.5	...
5	177.23	30	20	0	0.8	...	1	0	...
7	82.43	13	12	0.4	0.6	...	0.5	0	...
...

Como se puede observar, la distribución de los datos generados es desproporcionada. Las columnas Tiempo Invertido, Número de Sesiones y Días Activos son resultado de una suma, por lo que pueden aumentar sin fin. Por su parte, las columnas correspondientes a indicadores CD y SB tienen como valor máximo 1. Para que esta distribución de datos no produzca sesgos y resultados incorrectos en el proceso de clusterización se aplicó el algoritmo de la Normalización L2. La tabla de datos normalizada tiene la siguiente estructura:

Tabla 14

Indicadores Normalizados de Tiempo Invertido, Profundidad Cognitiva (CD) y Amplitud Social (SB)

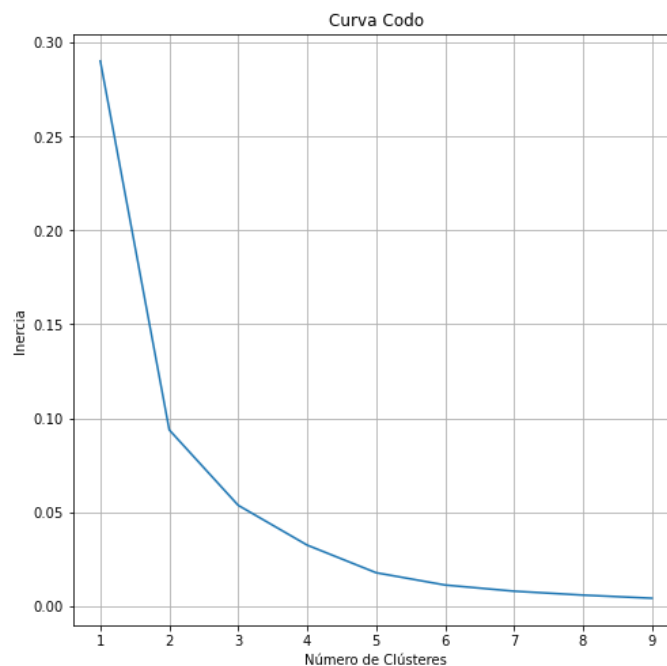
Usuario	Tiempo invertido	Número sesiones	Días activo	Recurso 1 (CD)	Recurso 2 (CD)	...	Recurso1 (SB)	Recurso 2 (SB)	...
4	0.91	0.91	0.73	1	0.5	...	0	0.5	...
5	0.56	0.53	0.31	0	0.8	...	1	0	...
7	0.25	0.21	0.15	0.4	0.6	...	0.5	0	...
...

Una vez los datos están normalizados, estimamos el k óptimo. En primer lugar, se utilizará el algoritmo de “Elbow Curve” mediante una implementación en el lenguaje de programación Python. El resultado del algoritmo es una gráfica en la que se debe identificar el punto de quiebre de una curva descendente (Curva del Codo), punto en el cual se ubica el número óptimo para k (Figura 38). La

métrica para el cálculo de las distancias elegida para el algoritmo es *inertia*, la cual se refiere a una suma de las distancias al cuadrado de las muestras con respecto al centro de agrupación más cercano.

Figura 38

Curva del Codo



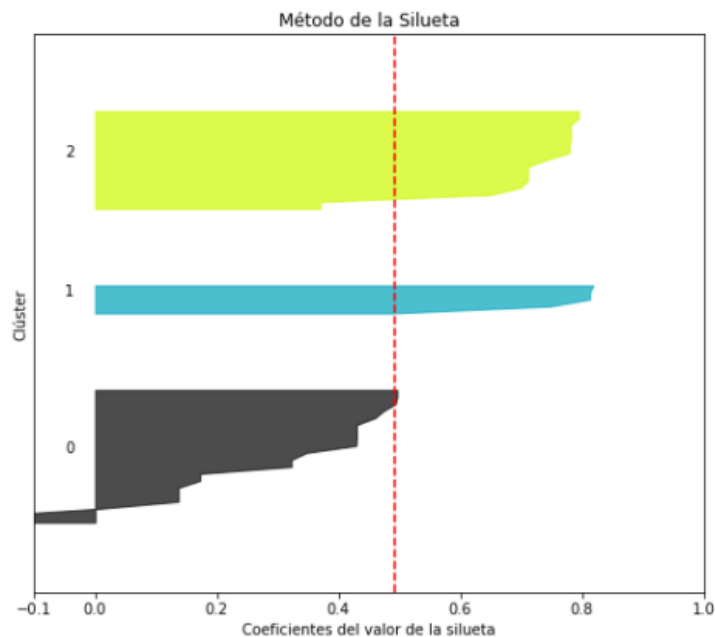
Nota. Elaboración Propia.

Como se observa, el primer punto de quiebre de la Curva del Codo se da en $k = 2$ y la curva presenta un decrecimiento menos considerable a partir de $k = 4$, por lo que cualquier valor para k entre 2 y 4 sería óptimo según este algoritmo. Para el caso de este estudio, se ha tomado el valor intermedio del rango, $k = 3$.

Para corroborar que la elección de un $k = 3$, se ha empleado el método de “Silhouette” o también conocido como el método de la Silueta. Para este método se utilizará el algoritmo de “Silhouette” mediante una implementación en el lenguaje de programación Python. Este método se utiliza para estudiar la distancia de separación entre los clústeres resultantes. El resultado de este algoritmo es un gráfico que muestra la medida de cuán cerca está cada punto en un clúster de los puntos en los clústeres vecinos. Esta medida tiene un rango de -1 a 1 . Donde 1 significa que los puntos están muy cerca de su propio clúster y lejos de otros clústeres. Mientras que, -1 indica que los puntos están cerca de los clústeres vecinos (Figura 39).

Figura 39

Gráficas del Método de la Silueta para $k = 3$



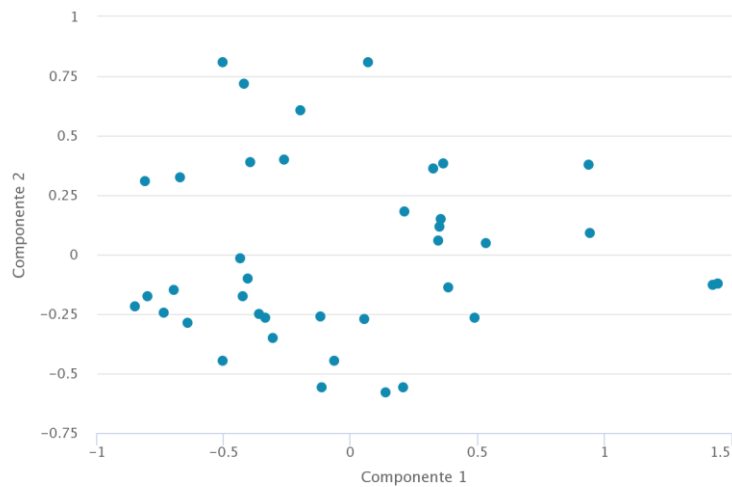
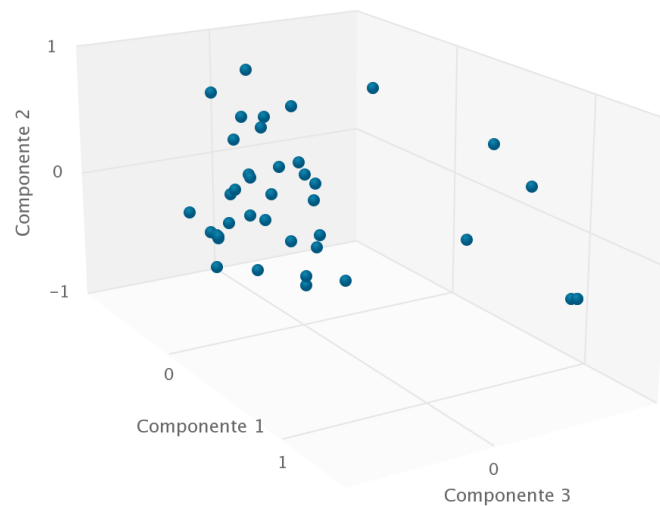
Nota. Elaboración Propia.

Entre el análisis de coeficientes para varios valores de k y las gráficas de silueta, se optó por un $k = 3$ debido a la distribución de datos entre cada cluster. Como se puede observar en la Figura 39, para $k = 3$ el coeficiente de la silueta se aproxima al valor de 1 . Esto indica que los puntos están cerca de su propio clúster y alejados de otros.



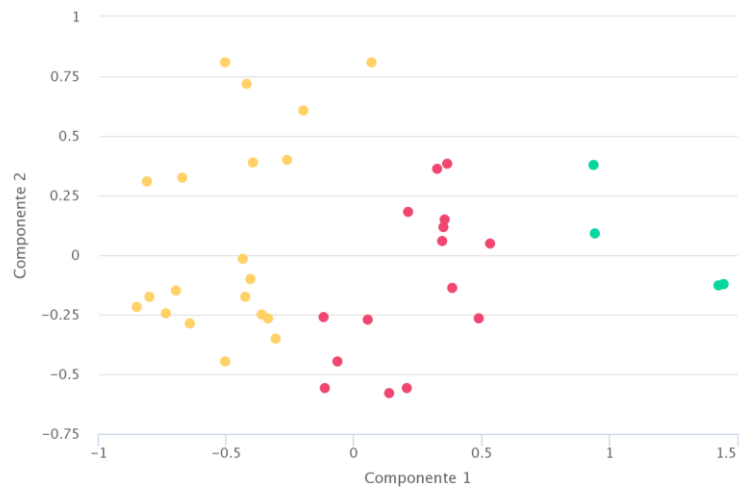
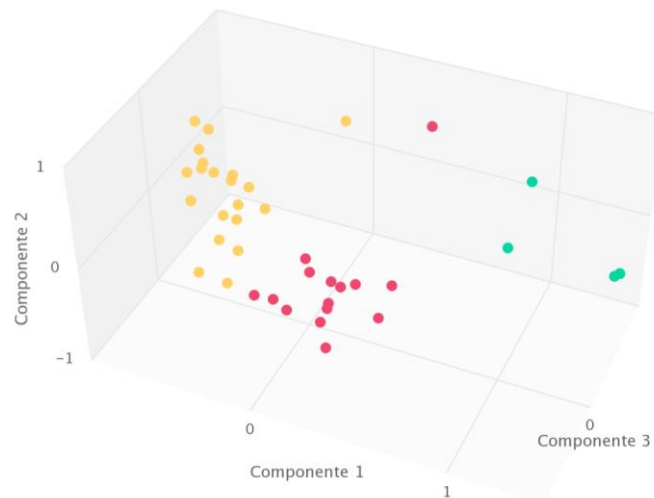
Los análisis previos nos dejan con la conclusión de que el número de Clusters óptimo para este conjunto de datos es $k = 3$. Una vez definido el valor de k , el siguiente paso es realizar el proceso de clusterización como tal.

En primer lugar, se va a realizar una gráfica de los datos para tener una primera vista de su distribución. Para esto, es necesario realizar una reducción de la dimensionalidad de los datos ya que, para el ejemplo propuesto, el conjunto de datos cuenta con un total de 148 columnas. La reducción realizada será a dos y tres componentes, con el fin de tener una visualización de los datos en dos y tres dimensiones (Figura 40).

Figura 40*Gráfico de Distribución Datos con Reducción de Componentes**a) Reducción a dos Componentes**b) Reducción a tres Componentes*

Nota. Elaboración Propia.

Luego de haber aplicado el algoritmo de K-Means a los datos de ejemplo, el resultado es una columna adicional generada por el algoritmo que especifica a qué clúster pertenece cada registro del conjunto de datos. Con estos datos se han realizado gráficas en dos y tres dimensiones para visualizar la distribución de los datos en cada clúster (Figura 41).

Figura 41*Gráfico de Distribución Datos Agrupados en tres Clústeres**c) Reducción a dos Componentes**d) Reducción a tres Componentes*

Nota. Elaboración Propia.

Como se puede observar en los gráficos, existen dos clústeres con la mayor cantidad de datos, mientras que un clúster tiene la menor cantidad de datos. Para realizar un análisis de los tipos de estudiantes de cada clúster se ha tomado en cuenta las variables Tiempo Invertido, Número de Sesiones y Días Activos. Se hizo la elección de estas variables ya que son los indicadores más

representativos y analizables para diferenciar a un estudiante. Con ellas, se caracterizó a cada clúster generador por el algoritmo (Tabla 15).

Tabla 15

indicadores más Relevantes que Caracterizan los Clústeres

Clúster	Promedio días activo	Promedio Tiempo Invertido (min)	Promedio Número Sesiones
0	17.40	397.55	34.06
1	22.25	559.10	42.50
2	4.00	48.79	6.40

Como se puede observar, las variables escogidas permiten caracterizar de buena manera a cada clúster. Se puede distinguir que los estudiantes pertenecientes al clúster 1 son los que en promedio, más tiempo han invertido en la plataforma y teniendo una mayor cantidad de sesiones y días activos en la plataforma. Estos estudiantes serían los que están más comprometidos con el curso (tiempo invertido e interacciones con material didáctico y actividades). Basado en los mismos parámetros, los estudiantes del clúster 0 también presentan valores considerablemente buenos para las variables elegidas. Esto significa que, si bien no son los más comprometidos con el curso, tienen un buen ritmo de avance y se les podría motivar a que aumenten su compromiso. Finalmente, los estudiantes que pertenecen al clúster 2 son los que peores valores tienen para las variables elegidas. Estos estudiantes son los que estarían en riesgo de deserción en el curso, dado que presentan la menor cantidad de interacciones y tiempo invertido en el curso. Además, se pudo confirmar esta afirmación ya que dos estudiantes que no presentaron interacción alguna con el curso terminaron agrupados dentro del clúster 2.

6.3.6. Métodos de división del tiempo

Gran parte de los estudios que han utilizado técnicas Machine Learning para predecir la deserción de los estudiantes en cursos online se han realizado con datos de cursos ya concluidos (Moreno Marcos, y otros, 2020); (Mwalumbwe & Mtebe, 2017); (Xing & Du, 2018). A pesar de que estas investigaciones han sido bastante beneficiosas, su utilidad para detectar estudiantes en riesgo de deserción a tiempo para hacer una intervención oportuna es limitada. Esto se da por la limitante de no poder clasificar al estudiante de manera clara durante el curso. Una posible solución podría ser utilizar los datos de cursos anteriores de estudiantes que aprobaron o reprobaron el curso como entrada a algún algoritmo de aprendizaje supervisado para entrenar un modelo predictivo que ayude



a identificar a los estudiantes en riesgo de deserción del nuevo curso, sin embargo, en la mayoría de las ocasiones no se cuentan con datos históricos que puedan servir como conjunto de entrenamiento.

Al no poder contar con datos históricos etiquetados, no es posible utilizar algoritmos de aprendizaje supervisado para predecir la deserción estudiantil. Por ello, el enfoque abordado en este estudio contempla el uso de un algoritmo de Machine Learning no supervisado (clusterización) para caracterizar a los estudiantes en base a su comportamiento en un determinado curso. De esta manera, ya no es necesario contar con datos históricos para realizar una predicción de la deserción. De esta manera, se tiene un método que puede agrupar a los estudiantes según su compromiso con el curso y se puede aplicar en cualquier momento durante el desarrollo del curso. Además, es necesario definir un intervalo de tiempo en el que se va a realizar la predicción de manera iterativa, aprovechando el hecho de que el transcurso del curso genera más datos que son útiles para mejorar las predicciones.

Para el caso de la herramienta que se desarrolla en este estudio, el rango de tiempo empieza en la fecha de inicio del curso y termina en la fecha de finalización. En este rango, la clusterización de los estudiantes se la realiza los viernes de cada semana, de manera que al finalizar una semana de trabajo el docente pueda evaluar a los estudiantes que están en riesgo de deserción y pueda elaborar material educativo adicional para la siguiente semana que motive a los estudiantes a mejorar su rendimiento académico. Además, con la herramienta el docente puede hacer un seguimiento de algunos indicadores (sesiones de estudio, interacciones con recursos, progreso del curso, calificaciones, tiempo invertido) de manera individual para cada estudiante, lo que le permite generar un perfil más específico del estudiante.

6.4. Planificación y Diseño del Dashboard

Como ya se mencionó, las herramientas de visualización de datos educativos ya han sido señaladas por la literatura en Learning Analytics como una de las brechas que existen actualmente en el área (Einhardt, Aires Tavares, & Cechinel, 2016). Aunque en los últimos años se han desarrollado aplicaciones de dashboard para apoyar el proceso de enseñanza-aprendizaje (Ginda, Suri, Bueckle, & Börner, 2016), la gran mayoría de ellas no están disponibles para descarga gratuita y uso general, ya que se han desarrollado para LMS pagados.

Antes de comenzar el desarrollo de la herramienta, se realizó un análisis de diferentes herramientas existentes con el objetivo de identificar posibles visualizaciones que resulten útiles en base a los datos previamente analizados. En este contexto se analizaron herramientas y plataformas que proporcionan visualizaciones similares que ayudaron a responder a las preguntas planteadas en la sección anterior. Las plataformas IntelliBoard y LearnerScript, además del plugin Note My



Progress fueron la mejor base de referencia para la investigación. En base al análisis realizado se han planteado prototipos de diferentes visualizaciones como se muestra a continuación.

Actividades de estudio

P_AE_1. Se requiere saber la cantidad de visitas realizadas por los estudiantes al curso. La información debe poder visualizarse por (día de la semana) / (hora del día) y por mes / (número de semana). También es necesario que se pueda visualizar el número de sesiones (Figura 42).



Figura 42

Prototipo para conocer la cantidad de visitas realizadas al curso por (día de la semana) / (hora del día) a modo de mapa de calor



Nota. Elaboración Propia.



P_AE_2. Se requiere visualizar un reporte con el número de visitas al curso de los estudiantes, la cantidad de tiempo han empleado en el curso y el progreso en el curso basado en las actividades realizadas y recursos revisados (Figura 43).

Figura 43

Prototipo para visualizar un reporte de visitas, tiempo y progreso de los estudiantes en el curso a modo de tabla interactiva

PREGUNTA DOCENTE 3: Visitas al Curso y Tiempo en el Curso

AYUDA ?

Participación de los estudiantes

Participante	Progreso	Visitas	Tiempo empleado
Juan Martinez	50%	5	01:05:19
Jhon Doe	70%	15	03:43:10
Sofia Narvaez	23%	2	00:10:15
Jhon Doe	70%	15	03:43:10
Juan Martinez	50%	5	01:05:19
Jhon Doe	70%	15	03:43:10
Sofia Narvaez	23%	2	00:10:15
Jhon Doe	70%	15	03:43:10
Juan Martinez	50%	5	01:05:19
Jhon Doe	70%	15	03:43:10
Sofia Narvaez	23%	2	00:10:15
Jhon Doe	70%	15	03:43:10

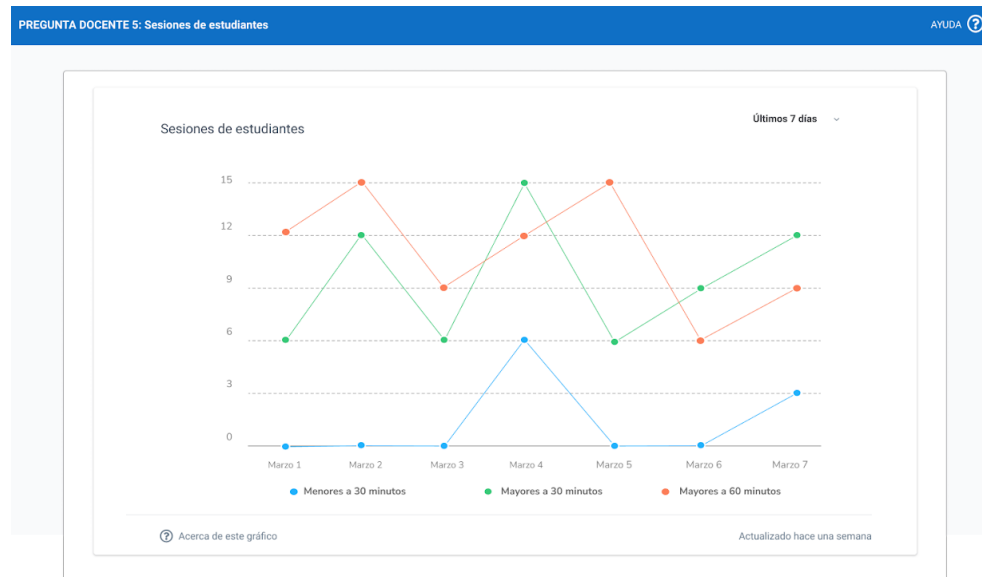
Rows per page : 5 1-10 of 100 < >

Nota. Elaboración Propia.

P_AE_3. Se desea conocer el número de sesiones menores a 30 minutos, mayores a 30 y menores a 60 minutos y mayores a 60 minutos en un curso (Figura 44).

Figura 44

Prototipo para visualizar la distribución de sesiones por diferentes rangos de tiempo mediante un gráfico de líneas



Nota. Elaboración Propia.

Seguimiento (tiempo)

P_ST_1. Se desea conocer el tiempo promedio invertido en el curso en comparación al tiempo promedio planificado por el docente a lo largo de las diferentes semanas planificadas en el curso (Figura 45).

Figura 45

Prototipo para conocer el tiempo promedio invertido a comparación del tiempo promedio planificado por el docente



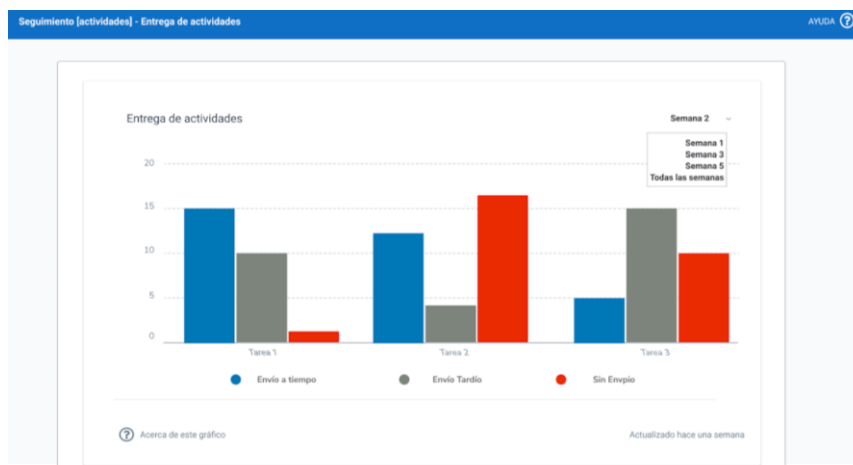
Nota. Elaboración Propia.

Seguimiento (actividades)

P_SA_1. Se desea conocer la cantidad de estudiantes que han enviado a tiempo, tarde o que no han enviado actividades de tipo assign en cada una de las diferentes semanas planificadas a lo largo del curso. Además, permitir la posibilidad de enviar un email a los estudiantes que han enviado, no han enviado o han enviado tarde una actividad de tipo **assign** (Figura 46).

Figura 46

Prototipo para conocer los Diferentes Estados de envío de Actividades para su Calificación

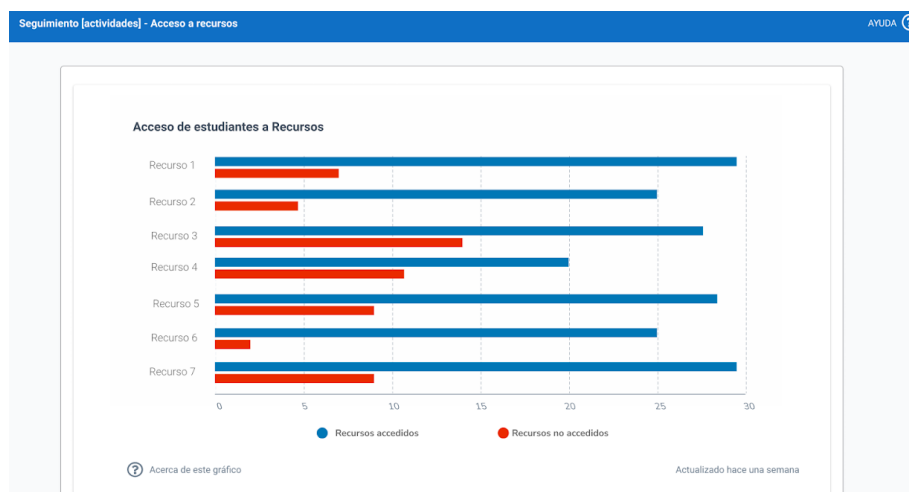


Nota. Elaboración Propia.

P_SA_2. Se desea conocer la cantidad de estudiantes que han accedido y no han accedido a los recursos y actividades (que el docente seleccione) durante las diferentes semanas de duración del curso. Además, permitir la posibilidad de enviar un email a los estudiantes que han accedido o no a dichos recursos y actividades (Figura 47).

Figura 47

Prototipo para conocer la Distribución de Acceso a los Recursos por parte de los Estudiantes



Nota. Elaboración Propia.



Seguimiento (para evaluaciones)

P_SE_1. Se desea conocer el promedio general de calificaciones de cada una de las actividades evaluables del curso. Al seleccionar una actividad específica mostrar la distribución de calificaciones con el promedio, la nota máxima y la nota mínima para dicha actividad (Figura 48).

Figura 48

Prototipo para conocer el Promedio de Calificaciones de las Diferentes Actividades Calificables

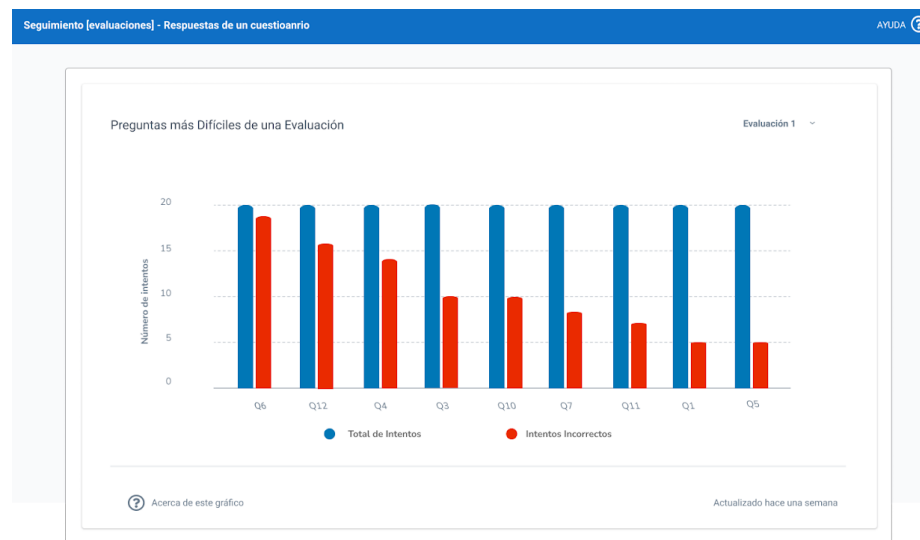


Nota. Elaboración Propia.

P_SE_2. Se desea conocer la cantidad de total de intentos, intentos parcialmente correctos, intentos incorrectos, intentos correctos, preguntas en blanco y sin calificar en cada una de las preguntas de una evaluación de todos los intentos realizados por los estudiantes

Figura 49

Prototipo para conocer la Distribución de Intentos para cada una de las Evaluaciones



Nota. Elaboración Propia.

6.5. Implementación del Dashboard

Este proyecto, como se comenta en el [Capítulo 1](#), consiste en el desarrollo de un plugin para Moodle. Por consiguiente, el objetivo final que se persigue con la elaboración del trabajo es su integración en esta plataforma. Para llegar a este objetivo, se siguieron una serie de pautas recomendadas desde la comunidad de desarrolladores de Moodle. Algunas de estas normas son estudiar la documentación que proporciona la plataforma en forma de tutoriales y guías o seguir el manual de estilo de código, donde se establecen las reglas generales para el tratamiento de ficheros, funciones y variables, la estructura de la base de datos y las normas de seguridad.

6.5.1. Descripción general

La herramienta de visualización en adelante FlipMyLearning (FML) fue construida para la plataforma Moodle 3.9+, es instalable y escalable, ya que está desarrollado bajo las directrices arquitectónicas modulares de la plataforma, lo cual le permite ser desinstalado e instalado en cualquier



momento y no contiene dependencias de otros módulos. Su objetivo es mostrar información, de manera gráfica e intuitiva, sobre el proceso de aprendizaje de los estudiantes en un curso de Moodle.

La herramienta de visualización permite la recopilación, preparación y transformación de los datos, así como el análisis de éstos a través de un proceso automático e intuitivo para obtener resultados que facilite a los docentes la comprensión del proceso de aprendizaje de sus estudiantes, basado en su interacción con los recursos del curso. Esto conlleva a presentar gráficos que muestran comportamientos que siguen los usuarios en una sesión, generados a partir de los datos obtenidos de la interacción con los diversos recursos propuestos por el docente.

Se aprovechan las características del entorno donde se ejecutará el plugin para crear una interfaz que permita a los docentes (de diversas áreas) la capacidad de interactuar con el mismo, sin requerir de conocimiento previo del desarrollo del plugin. Los resultados obtenidos permiten a los docentes evidenciar el progreso de cada estudiante en base a actividades calificables del curso, acceso a recursos, cuestionarios, tiempo empleado y sesiones de estudio. Este plugin facilita la tarea de los docentes al analizar grandes volúmenes de información mediante la analítica del aprendizaje y permite una mejora en la toma de decisiones.

6.5.2. Arquitectura

La arquitectura implementada se denomina Modelo Vista Controlador (*MVC – Model View and Controller en inglés*), la cual contempla una estructura de tres capas, separando la interfaz, el comportamiento y los datos (Aguilar, 2019). La arquitectura MVC es muy utilizada en el desarrollo Web, pues permite la representación de un mismo conjunto de datos de forma diferente, lo que se denomina una Vista, y asocia cada Vista a un comportamiento particular de la aplicación, denominado Modelo. De esta forma, cuando los datos se modifican en una de las Vistas, todas las representaciones también se actualizan.

Cuando el usuario interactúa con una de las vistas, el controlador detecta el cambio y solicita los datos mediante JavaScript Asíncrono y XML (*AJAX – Asynchronous JavaScript and XML en inglés*) a alguno de los modelos ya definidos que proporciona Moodle mediante su Interfaz de Programación de Aplicaciones (*API – Application Programming Interface en inglés*) de Data Manipulation. La Figura 50 muestra el flujo de información del plugin. Se puede observar en la imagen, que el flujo de información inicia en la capa de datos, en donde el plugin utiliza funciones para recolectar información desde la base de datos. Posteriormente, una vez hecha la consulta, se procesan los datos en los ficheros de preprocesamiento y construcción de información, para

finalmente ser enviados en formato de Notación de Objetos de JavaScript (*JSON – JavaScript Object Notation en inglés*) a la capa de aplicación o interfaz de usuario.

Figura 50

Arquitectura General



Nota. Elaboración Propia.

En resumen, la primera de las capas está representada por todos los archivos lógicos (funciones de conexión a la base de datos) que consultan los datos guardados en tablas de la plataforma. La segunda está formada por todos los ficheros que trabajan para construir estructuras de datos a partir de los datos extraídos. Finalmente, la tercera y última parte está formada por todas las bibliotecas visuales que se utilizan para presentar los datos visuales al usuario.

6.5.3. Tecnologías involucradas

Para desarrollar la herramienta FlipMyLearning se han seleccionado tecnologías estándares que sean de libre distribución e independientes de cualquier plataforma de ejecución, lo que garantiza su portabilidad, disponibilidad, compatibilidad y eficiencia (Einhardt, Aires Tavares, & Cechinel, 2016). A continuación, se listan las tecnologías que fueron usadas para la creación de la herramienta de visualización.

- **Framework Vue JS:** es un framework progresivo para construir interfaces de usuario. Un framework es un conjunto de herramientas para el desarrollo que brindan cierta metodología de trabajo. Por progresivo se entiende que Vue está diseñado desde cero para ser utilizado incrementalmente junto al proyecto. La librería central está enfocada solo en la capa de visualización, y es fácil de utilizar e integrar con otras librerías o proyectos existentes (Vue.js Team, 2020).



Se seleccionó este framework debido a que se caracteriza por trabajar con el patrón denominado MVVM que significa Modelo Vista - Vista Modelo que simplifica el desarrollo de interfaces impulsado por eventos. Además, es un framework que puede ser integrado con cualquier lenguaje de programación o incluso framework. Entre sus principales características están: ser reactivo, sencillo de aprender, ligero, tienen un estilo de desarrollo orientado a componentes y gran soporte por parte de la comunidad.

- **Git y Github:** por un lado, git es un software de control de versiones para desarrolladores. El control de versiones se refiere al proceso de guardar diferentes archivos o “versiones” a lo largo de las diferentes etapas de un proyecto. Por otro lado, GitHub facilita la colaboración con git. Es una plataforma que puede mantener repositorios de código en almacenamiento basado en la nube para que varios desarrolladores puedan trabajar en un solo proyecto y ver las ediciones de cada uno en tiempo real. Además, los repositorios de GitHub están disponibles públicamente. Los desarrolladores de todo el mundo pueden interactuar y contribuir al código de los demás para modificarlo o mejorarlo, lo que se conoce como “codificación social”.

Se seleccionó este software ya que Moodle comparte su código fuente a través de Github², de manera que con Git y Github es posible clonar el código fuente de Moodle para trabajar de manera local en un servidor.

- **PHP:** para el desarrollo de componentes de software, funciones y bibliotecas. Es el lenguaje fundamental de la plataforma Moodle y su intérprete puede integrarse de manera eficiente en el servidor Web Apache.
- **HighCharts:** librería escrita en Javascript que permite la creación de gráficas. La librería ofrece un método fácil e interactivo para insertar gráficas en un sitio o aplicación web. Es una librería compatible con todos los navegadores modernos. No es comercial y todas las características pueden ser personalizadas permitiendo una gran flexibilidad (Highcharts Team, 2020).

Se seleccionó esta librería ya que como está escrita en JavaScript, posee una buena integración con código PHP, además tiene una extensa variedad de gráficos disponibles para su uso y personalización. Únicamente es necesario dar a los datos el formato adecuado para poder mostrar las visualizaciones disponibles, sin mencionar que cuenta con una amplia documentación que la respalda.

² Repositorio del código fuente de Moodle <https://github.com/moodle/moodle>



- **Lenguaje HTML:** (*HyperText Markup Language en inglés*) para la elaboración de las páginas Web que conforman la interfaz visible del componente. Establece la estructura y contenido de la página.
- **Lenguaje CSS:** (*Cascade Style Sheets en inglés*) para definir la presentación de una página Web en términos de un formato y tipografía específica, independiente de la implementación. A los archivos CSS se les denomina hojas de estilo.
- **PHP-ML – Machine Learning library for PHP:** php-ml es una biblioteca de aprendizaje automático escrita en PHP. Posee algoritmos de aprendizaje automático, validación cruzada, redes neuronales, preprocesamiento, clasificación, agrupamiento, extracción de características entre otros (Arkadiusz, 2020).

Se seleccionó esta librería ya que como está escrita en PhP, contiene una buena integración con Moodle, además posee algoritmos necesarios para aplicar técnicas de minería de datos como clasificación, regresión, clustering, entre otras. Posee una fácil integración gracias a su licencia MIT, es decir, impone muy pocas limitaciones en la reutilización y el código se encuentra disponible en su repositorio oficial de GitHub³.

- **Lenguaje JavaScript:** para manipular dinámicamente los distintos elementos de la interfaz, lo que permite establecer nuevos estilos de interacción como AJAX. Generalmente los archivos JavaScript constituyen componentes interactivos.
- **MySQL:** Sistema de Gestión de Bases de Datos Relacionales (*RDBMS – Relational Database Management System en inglés*) con un modelo cliente-servidor. RDBMS es un software o servicio utilizado para crear y administrar bases de datos basadas en un modelo relacional (MySQL Team, 2020).

6.5.4. Proceso de instalación

FlipMyLearning se instala en la carpeta local del directorio de instalación de la plataforma Moodle. Es decir, es un tipo de plugin local de la lista de tipos de plugins de Moodle (Moodle, 2020). Este tipo de plugin es la forma recomendada de agregar nuevas funcionalidades a la plataforma y proporciona una manera sencilla de instalación. Debido a la modularidad de Moodle, únicamente es necesario el directorio con la funcionalidad específica (Moodle, 2020). Para ejecutar una instalación de FlipMyLearning es necesario poseer una cuenta de administrador del sitio y seguir los siguientes pasos:

³ Repositorio del código fuente de php-ml <https://github.com/jorgecasas/php-ml>

- 151



Figura 53

fliplearning en el directorio local de Moodle

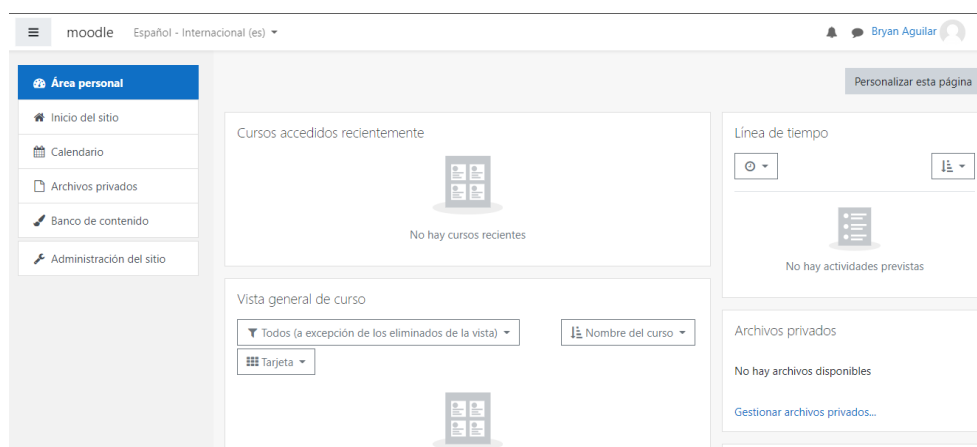


Nota. Elaboración Propia.

- Una vez el directorio **fliplearning** se encuentre en el directorio local, es necesario volver al sitio de Moodle y mantener una sesión iniciada como administrador. Es recomendable posicionarse en la sección de Área Personal del sitio como se muestra en la Figura 54.

Figura 54

Sesión de administrador iniciada en la sección de Área Personal del sitio



Nota. Página inicial de Moodle luego de iniciar sesión como administrador. Elaboración Propia.

- Ubicado en la sección de Área Personal de Moodle únicamente será necesario recargar el sitio para ejecutar la instalación. Al recargar el sitio se obtiene una pantalla de instalación como se muestra en la Figura 55.



Figura 55

Comprobación de requisitos y confirmación de instalación para FlipMyLearning

Información de la versión

Moodle 3.9.2+ (Build: 20201030)

Si desea información sobre esta versión de Moodle, por favor vea [Release Notes](#)

Comprobaciones del servidor

Nombre	Información	Informe	Plugin	Estado
moodle		versión 3.5 es obligatoria y está ejecutando 3.9.2+ (Build: 20201030)		OK
unicode		debe estar instalado/activado		OK
database	mysql (8.0.21)	versión 5.6 es obligatoria y está ejecutando 8.0.21		OK
php		versión 7.2.0 es obligatoria y está ejecutando 7.3.21		OK
pcreunicode		debería estar instalado y activado para conseguir los mejores resultados		OK
php_extension	iconv	debe estar instalado/activado		OK
php_extension	mbstring	debe estar instalado/activado		OK
php_extension	curl	debe estar instalado/activado		OK
php_setting	memory_limit	detectado ajuste recomendado		OK
php_setting	file_uploads	detectado ajuste recomendado		OK
php_setting	opcache.enable	detectado ajuste recomendado		OK

Otras comprobaciones

Información	Informe	Plugin	Estado
site_not_https	<p>Si esta comprobación falla, ello indica un problema potencial</p> <p>Se ha detectado que su sitio no se comunica a través de HTTPS. Se recomienda migrar su sitio a HTTPS para incrementar la seguridad y mejorar la integración con otros sistemas.</p>		Revisar
mysql_full_unicode_support	<p>Si esta comprobación falla, ello indica un problema potencial</p> <p>La configuración actual de MySQL o MariaDB está usando 'utf8'. Este set de caracteres no soporta caracteres de 4 byte los cuales incluyen algunos emoji. Si intenta usar esos caracteres finalizará con un error al actualizar una fila, y toda la información que esté siendo enviada a la base de datos se perderá. Por favor considere cambiar la configuración a 'utf8mb4'. Mire la documentación para más información.</p>		Revisar

Su entorno de servidor cumple todos los requerimientos mínimos.

NOTA: Moodle intentará descargar actualizaciones para sus paquetes de idioma durante la actualización.

[Continuar](#)

Nota. Página de comprobaciones necesarias de software previo a instalar un plugin de Moodle. Elaboración Propia.

- En dicha pantalla únicamente es necesario dar clic en el botón continuar. Posteriormente se muestra una pantalla que indica que se requiere actualizar la base de datos de Moodle (Figura 56) ya que FlipMyLearning crea y usa tablas de datos independientes de las tablas de datos de Moodle.



Figura 56

Confirmación de Actualización de la Base de Datos de Moodle

Plugins que requieren su atención

[Plugins que requieren su atención](#) 1 [Todos los plugins](#) 412

Nombre de la extensión / Directorio	Versión actual	Nueva versión	Requiere	Origen / Estado
Extensiones (plugins) locales				
Flip My Learning /local/fliplearning		2020111018	• Moodle 2015111601	Adicional Para instalarse
Recargar				
Actualizar base de datos Moodle ahora				

Nota. Página de confirmación de actualización de la base de datos previo a instalar un plugin de Moodle.
Elaboración Propia.

- Una vez finalizada la actualización de la base de datos de Moodle. El sitio mostrará una pantalla indicando que el plugin FlipMyLearning ha sido instalado de manera exitosa. Para finalizar este proceso es necesario dar clic en la opción *Continuar*.

Figura 57

Confirmación de Instalación Exitosa

Actualizando la versión

local_fliplearning

Éxito

Continuar

Nota. Elaboración Propia.

- Una vez finalizado el proceso de instalación es posible dirigirse a la página de un curso específico y configurar FlipMyLearning como se indica en la siguiente sección.

6.5.5. Funcionalidades

A continuación, se describen todas las funcionalidades que tiene FlipMyLearning para Moodle, cuando se realiza el análisis de un curso. La herramienta toma en consideración el tiempo que los estudiantes invierten en el curso, interacciones con recursos, entrega de tareas, sesiones de estudio y rendimiento en evaluaciones como elementos esenciales para el análisis de la información.

Cuando la herramienta es instalada por primera vez en el servidor de la plataforma Moodle el docente o administrador del sitio puede acceder a ella desde la parte superior derecha de la ventana de un curso de Moodle (Figura 58). La herramienta puede ser identificada mediante un ícono como se muestra en la Figura 59. Para que la herramienta muestre las visualizaciones correspondientes es necesario realizar una configuración de semanas (Figura 60) como se explica en la siguiente sección.

Figura 58

Ubicación de FlipMyLearning en Moodle



Nota. Para acceder a FlipMyLearning es necesario posicionarse sobre un curso específico. Elaboración Propia.



Figura 59

Ícono de la Herramienta FlipMyLearning



Nota. Tomada de la Librería de iconos Fontawesome (<https://fontawesome.com/icons/chart-pie>).

Figura 60

Opción Inicial de Configuración de Semanas de Estudio



Nota. Para acceder a las funcionalidades de FlipMyLearning es necesario configurar las semanas de estudio.
Elaboración Propia.

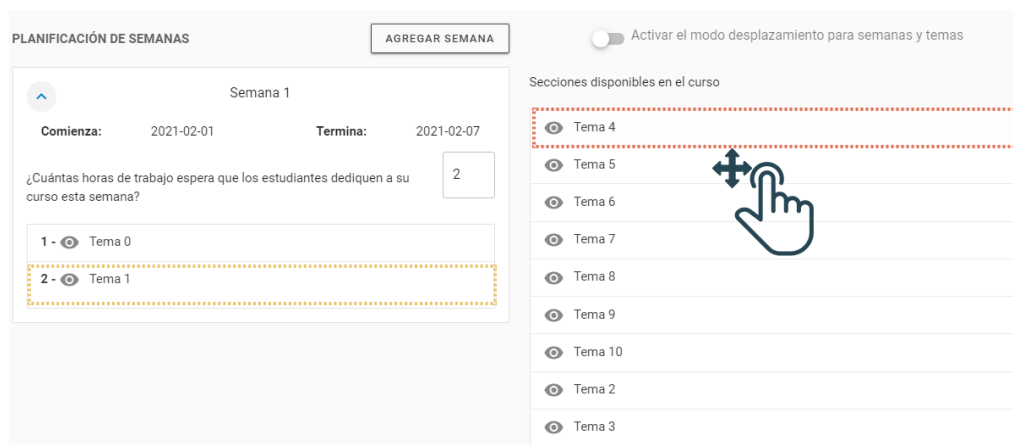
6.5.5.1. Configuración de semanas de estudio

Para el uso de la herramienta FlipMyLearning es necesario realizar una configuración inicial basada en semanas a lo largo del curso. En primer lugar, el docente debe definir la cantidad de semanas que durará su curso y agregarlas en la sección *Planificación de Semanas*. En segundo lugar, debe establecer las horas de trabajo que el estudiante debería invertir en cada una de las semanas a lo largo del curso. Por último, debe asignar una o varias secciones de su curso a cada semana que definió en el primer paso. Para asignar las secciones a las semanas el docente tiene la posibilidad de

arrastrarlas de la sección *Secciones disponibles en el curso* al panel de cada semana, como se muestra en la Figura 61. Esta configuración semanal puede realizarse desde la fecha de inicio del curso (creación del curso) en adelante, es decir FlipMyLearning calcula de manera automática las fechas de inicio y fin de la siguiente semana del curso.

Figura 61

Configuración de Semanas de Estudio



Nota. Adaptada de FlipMyLearning (Configurar Semanas).

Adicionalmente, esta configuración proporciona mensajes de advertencia en el caso de que el usuario genera una configuración incorrecta en las fechas. Para registrar una configuración adecuada de semanas de estudio se deben cumplir las siguientes condiciones:

- El curso debe poseer al menos un estudiante matriculado.
- La fecha actual debe ser mayor a la fecha de inicio de la primera semana configurada.
- Las semanas configuradas poseen al menos una sección.

En el caso de que al menos una de estas condiciones no se cumpla FlipMyLearning indicará la advertencia correspondiente como se muestra en la Figura 62.

Una vez realizada la configuración de las semanas de estudio es posible acceder a las secciones de la herramienta. FlipMyLearning cuenta con una vista para el docente y una vista para el estudiante. La vista para el docente cuenta con las secciones de: Indicadores Generales, Sesiones de Estudio, Seguimiento de Tareas, Seguimiento de Calificaciones, Seguimiento de Evaluaciones y

Deserción (Figura 63.a). Por el contrario, la vista para el estudiante cuenta con las secciones de: Indicadores Generales y Sesiones de Estudio (Figura 63.b). Ambas vistas dependen del rol del usuario que inicie sesión en la plataforma, es decir, si el docente o administrador inicia sesión únicamente verá las secciones del docente, por otro lado, si un estudiante inicia sesión únicamente tendrá acceso a las secciones del estudiante.

Figura 62

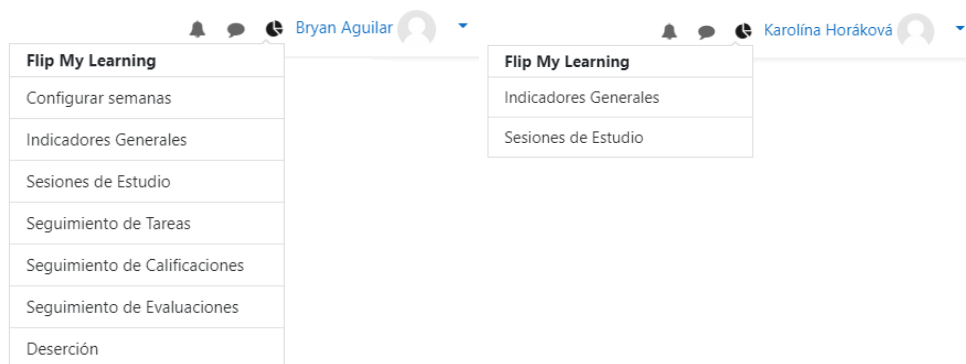
Advertencias de FlipMyLearning al no Cumplir una de las Condiciones de Configuración establecidas



Nota. Adaptada de FlipMyLearning (Configurar Semanas).

Figura 63

Secciones Disponibles de FlipMyLearning después de la Configuración de Semanas



a) Secciones disponibles para el docente

b) Secciones disponibles para el estudiante

Nota. Tomada de FlipMyLearning (menú de opciones).

Todas las secciones mostradas en la Figura 63 cuentan con una cabecera como se muestra en la Figura 64. Esta división se muestra de color azul en la parte superior e indica el nombre de la sección de FlipMyLearning y un botón de ayuda que muestra información referente a la visualización/nes mostradas en la sección.

Figura 64

Sección de Ayuda para Secciones de FlipMyLearning

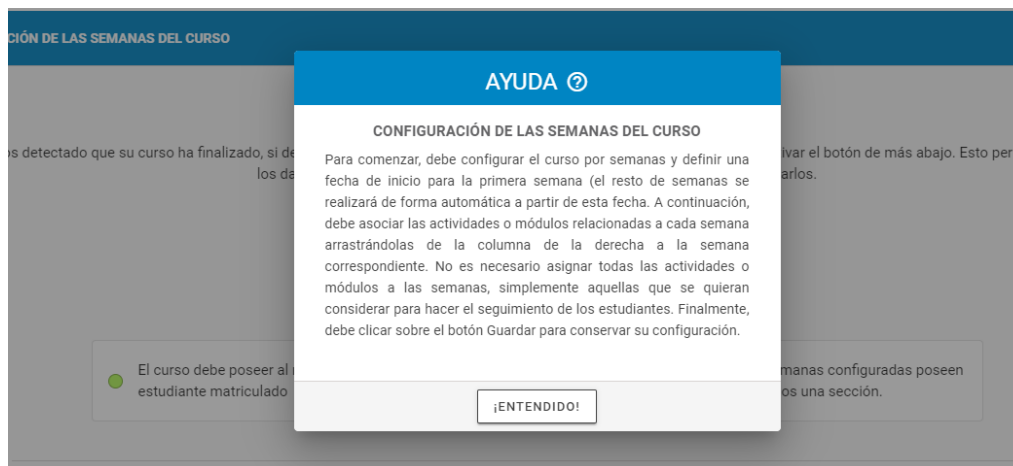


Nota. Tomada de FlipMyLearning (Configurar semanas de la vista del Docente).

Al presionar el botón de ayuda se muestra un modal (complemento para agregar diálogos, notificaciones y alertas al usuario con contenido personalizado) con ayuda al usuario sobre dicha sección como se muestra en la Figura 65. De esta manera el usuario podrá conocer de mejor manera la funcionalidad específica de una sección con el objetivo de mejorar su comprensión y uso.

Figura 65

Modal de Ayuda para las Diferentes Secciones de FlipMyLearning



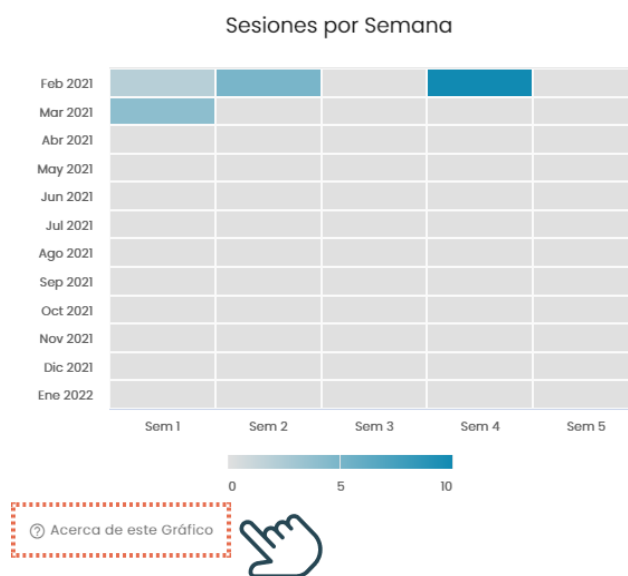
Nota. Tomada de FlipMyLearning (Configurar semanas de la vista del Docente).

Además de lo anterior, las visualizaciones de cada sección contienen en su parte inferior una característica llamada *Acerca de este gráfico* (Figura 66). Al presionar el botón de ayuda se muestra un modal con ayuda al usuario sobre dicha sección como se muestra en la Figura 65. De esta manera

el usuario podrá conocer de mejor manera la funcionalidad específica de una visualización. Adicionalmente, todas las visualizaciones tienen un tooltip (herramienta de ayuda visual, funciona al situar el cursor sobre algún elemento gráfico) que muestra información adicional del gráfico al docente o al estudiante.

Figura 66

Función Acerca de este gráfico para las Diferentes Visualizaciones de FlipMyLearning



Nota. Tomada de FlipMyLearning (Sesiones de Estudio de la vista del Docente).

6.5.5.2. Indicadores Generales (Vista Docente)

Esta sección contiene visualizaciones con indicadores generales relacionados a la configuración del curso, recursos asignados por semanas, sesiones de estudio y progreso de los estudiantes a lo largo del curso. Las visualizaciones de esta sección muestran los indicadores desde la fecha de inicio hasta la de finalización del curso (o hasta la fecha actual en caso de que el curso aún no ha terminado). Esta sección cuenta con cuatro visualizaciones distintas como se explican a continuación.

6.5.5.2.1. Información del Curso

La primera visualización muestra indicadores generales del curso. Estos indicadores son: la fecha de inicio del curso, la fecha de finalización del curso (en caso de estar configurada), el formato

del curso, la completitud de módulos (activada o desactivada), la cantidad de estudiantes inscritos (hasta la fecha de visualización), la cantidad de recursos y actividades que el docente ha puesto a disposición de los estudiantes, la cantidad de semanas de estudio configuradas y la calificación máxima del curso (basado en las ponderaciones máximas de actividades) (Figura 67).

Figura 67

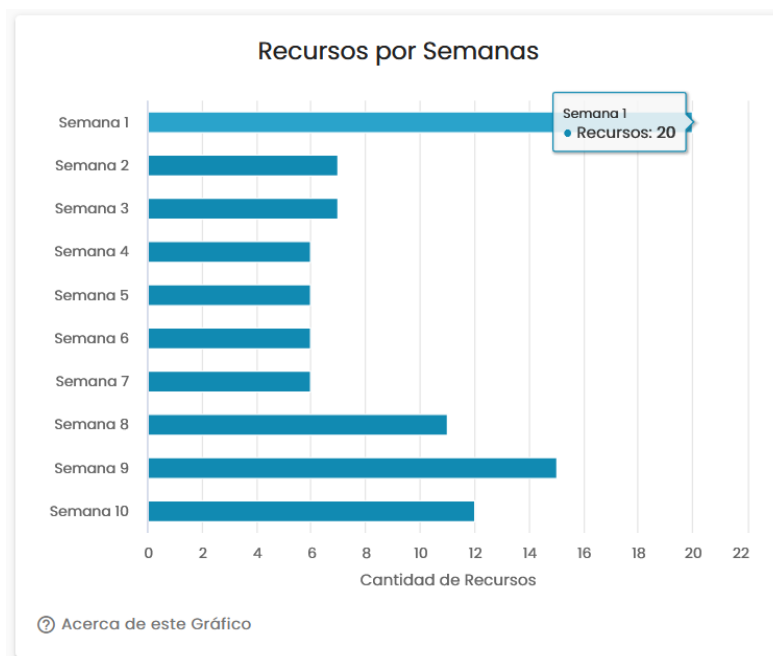
Indicadores Generales del Curso

Test Course 1			
INICIO Lunes, Noviembre 2 2020, 0:00	FIN Domingo, Febrero 14 2021, 23:59	FORMATO Tema	COMPLETITUD DE MÓDULOS Desactivado
58 Estudiantes	96 recursos	10 Semanas	2350 Calificación

Nota. Tomada de FlipMyLearning (Indicadores Generales de la vista del Docente).

6.5.5.2.2. Recursos por Semana

La segunda visualización se denomina *Recursos por Semana* (Figura 68). Este gráfico visualiza la cantidad de recursos de cada una de las secciones del curso asignadas a cada semana de estudio configurada en la sección *Configurar Semanas*. Si una semana tiene asignada dos o más secciones del curso, los recursos de dichas secciones se suman para el cálculo del total de recursos de una semana.

Figura 68*Recursos por Semanas de Estudio*

Nota. Tomada de FlipMyLearning (Indicadores Generales de la vista del Docente).

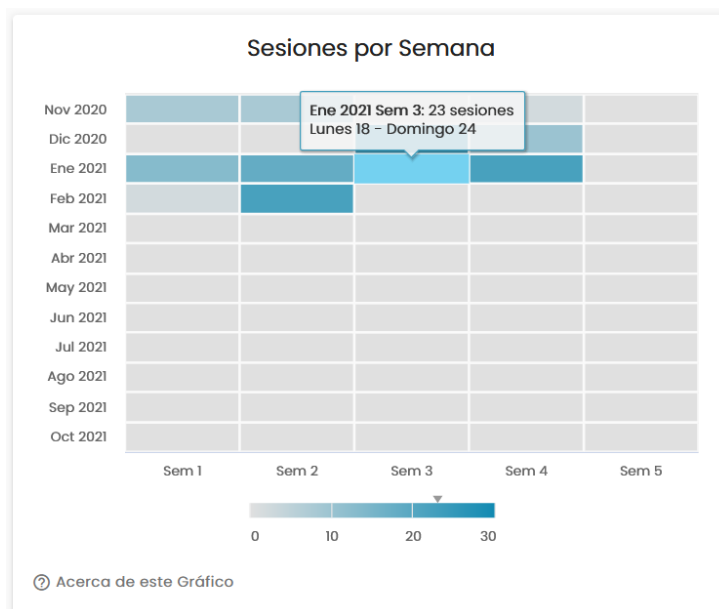
En el eje x del gráfico se encuentran el total de recursos y actividades de las secciones asignadas a cada semana configurada de FlipMyLearning. En el eje y se encuentran las semanas de estudio configuradas.

6.5.5.2.3. *Sesiones por Semana*

La tercera visualización se denomina *Sesiones por Semana*. Esta visualización permite evidenciar las sesiones de estudio por número de semana a lo largo de los diferentes meses del año en un mapa de calor (*heat map en inglés*) (Figura 69).

Figura 69

Mapa de calor de Sesiones por Semana



Nota. Mapa de calor para evidenciar las sesiones de estudio por semanas del mes. Tomada de FlipMyLearning (Indicadores Generales de la vista del Docente)

En el eje x se encuentran las semanas de cada uno de los meses. En el eje y del gráfico se encuentran los diferentes meses del año partiendo del mes de creación del curso. El gráfico resulta útil para el docente ya que permite visualizar de manera rápida en qué semanas del mes los estudiantes presentan mayor cantidad de sesiones al curso.

Las sesiones se identificaron en función de los accesos al curso y en base a un tiempo máximo de inactividad de 30 minutos entre dos interacciones de un mismo estudiante. Eso significa que cualquier interacción que se encuentren dentro de los 30 minutos se consideran parte de la misma sesión. De esta manera, si el tiempo entre dos interacciones es de más de 30 minutos, se consideran diferentes sesiones.

6.5.5.2.4. Progreso de los Estudiantes

La última visualización en esta sección es una tabla denominada *Progreso de los Estudiantes*. Esta tabla muestra la foto del perfil (en caso de estar disponible), nombres y apellidos, correo electrónico, *progreso en el curso*, sesiones de estudio y el tiempo total invertido de todos los estudiantes matriculados en el curso. Además, la tabla posee acciones como el ordenamiento de los



estudiantes por nombres, apellidos, progreso, sesiones o tiempo invertido. Búsqueda de un estudiante y paginación de registros para una mejor visualización. Esta tabla se puede observar en la Figura 70.

Figura 70

Tabla Interactiva para Evidenciar un Reporte de Participación de los Estudiantes

Progreso del los Estudiantes						
	Nombre	Apellidos	Correo	Progreso (%)	Sesiones	Tiempo Invertido
	秀英	王	tool_generator_000002@example.com	56 %	41	08:32:47
	Дарья	Новикова	tool_generator_000003@example.com	26 %	30	02:57:14
	Lukas	Meyer	edissonfernando2012@hotmail.com	16 %	13	01:22:25
	娜	周	tool_generator_000011@example.com	3 %	4	00:05:22
	Paul	Müller	tool_generator_000010@example.com	32 recursos finalizados de 96	4	00:04:59
	Ava	Garcia	edissonf.sigua@gmail.com	33 %	12	01:24:17
	Jakub	Černý	tool_generator_000006@example.com	6 %	8	00:09:30
	Matěj	Novák	tool_generator_000009@example.com	3 %	5	00:48:05
	Leonie	Hoffmann	tool_generator_000007@example.com	5 %	6	00:02:18
	Jayden	Jones	tool_generator_000008@example.com	4 %	3	00:02:52

Rows per page: 10 1-10 of 58

Acerca de esta Tabla

Nota. Tomada de FlipMyLearning (Indicadores Generales de la vista del Docente).

La información de esta tabla permite contestar la siguiente interrogante:

P_AE_2. Se requiere visualizar un reporte con el número de visitas al curso de los estudiantes, la cantidad de tiempo han empleado en el curso y el progreso en el curso basado en las actividades realizadas y recursos revisados.

El cálculo del progreso de un estudiante en el curso hace uso de la configuración de Moodle denominada Finalización de Actividad (*Activity Completion en inglés*). Esta configuración permite al docente configurar criterios de finalización dentro de la configuración de una actividad específica. Una casilla de verificación aparece junto a la actividad cuando el estudiante cumple el criterio. El criterio de finalización de la actividad depende del tipo de la actividad (los criterios varían dependiendo de los diferentes tipos de [actividades](#) disponibles en Moodle), pero los valores comunes para todas son: si la actividad ha sido vista, si el estudiante ha recibido una calificación o si el estudiante ha marcado manualmente como completa la actividad. En caso de que esta configuración se encuentre activada, FlipMyLearning calculará el progreso en el curso en base al número de



actividades completadas a lo largo de las semanas de estudio. Si alguna actividad no tiene activada dicha configuración, FlipMyLearning únicamente validará que dicha actividad haya sido vista por el estudiante para marcarla como completada. Con el objetivo de obtener mejores resultados en las visualizaciones de FlipMyLearning es necesario que el docente configure la Finalización de Actividad con criterios personalizados para las diferentes actividades propuestas en el curso a lo largo de las semanas de estudio.

6.5.5.3. Sesiones de estudio (Vista Docente)

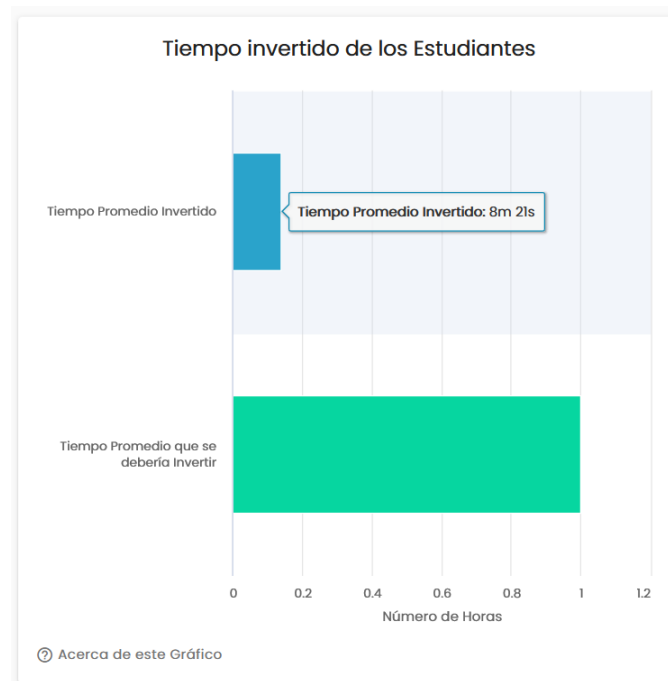
Esta sección contiene visualizaciones con indicadores relacionados a la actividad de los estudiantes en el curso medida en términos de sesiones realizadas, tiempo promedio invertido en el curso por semana y sesiones de estudio en intervalos de tiempo. Los datos presentados en esta sección varían dependiendo de la semana de estudio seleccionada.

6.5.5.3.1. Tiempo Invertido

La primera visualización muestra el tiempo promedio invertido por parte de los estudiantes en la semana seleccionada en comparación del tiempo promedio planificado por parte del docente en forma de un gráfico de barras horizontal (Figura 71).

Figura 71

Gráfico de barras para evidenciar el Tiempo promedio invertido por los Estudiantes a Comparación del Tiempo promedio planificado por el Docente a lo largo de una semana de estudio



Nota. Tomada de FlipMyLearning (Sesiones de Estudio de la vista del Docente)

Esta visualización permite responder la siguiente pregunta:

P_ST_1. Se desea conocer el tiempo promedio invertido en el curso en comparación al tiempo promedio planificado por el docente a lo largo de las diferentes semanas planificadas en el curso.

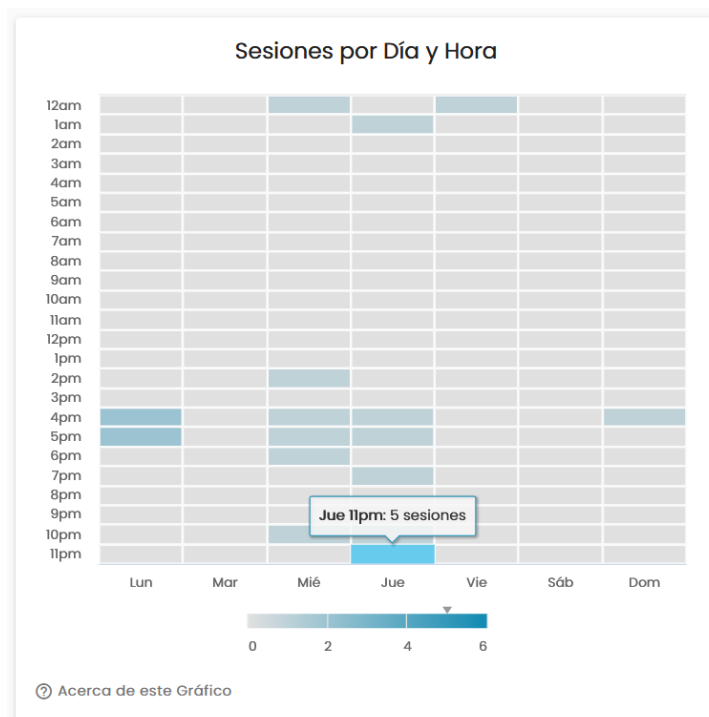
En el eje x del gráfico se encuentra el número de horas que el docente ha planificado para una semana específica y en el eje y se encuentran las etiquetas de tiempo promedio invertido y tiempo promedio que se debería invertir.

6.5.5.3.2. Sesiones por Día y Hora

La segunda visualización muestra un gráfico con las sesiones de estudio por día y hora de la semana seleccionada en un mapa de calor (Figura 72).

Figura 72

Mapa de calor de Sesiones por Día y Hora



Nota. Tomada de FlipMyLearning (Sesiones de Estudio de la vista del Docente).

Esta visualización ayuda a responder la siguiente pregunta:

P_AE_1. Se requiere saber la cantidad de visitas realizadas por los estudiantes al curso. La información debe poder visualizarse por (día de la semana) / (hora del día). También es necesario que se pueda visualizar el número de sesiones.

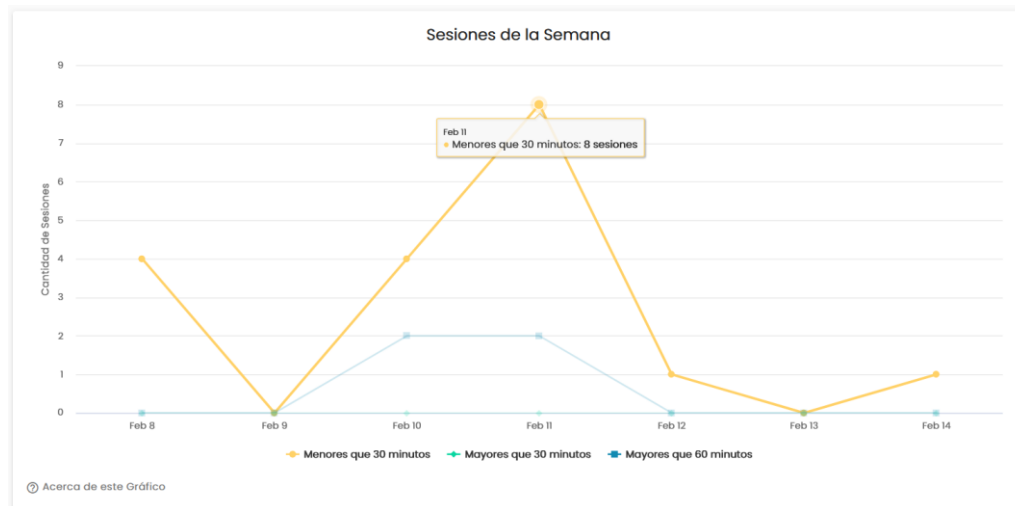
En el eje x del gráfico se encuentran los diferentes días de la semana y en el eje y se encuentran las horas del día empezando por las 12am y terminando a las 11pm o 23 horas.

6.5.5.3.3. Sesiones de la Semana

La tercera visualización muestra en un gráfico de líneas el número de sesiones realizadas en la semana clasificadas por su duración en rangos de tiempo: menores a 30 minutos, mayores a 30 minutos y mayores a 60 minutos (Figura 73).

Figura 73

Gráfico de Líneas para evidenciar la Cantidad de Sesiones en Diferentes intervalos de tiempo a lo largo de una semana de estudio



Nota. Tomada de FlipMyLearning (Sesiones de Estudio de la vista del Docente).

Esta visualización permite responder a la siguiente pregunta:

P_AE_3. Se desea conocer el número de sesiones menores a 30 minutos, mayores a 30 y menores a 60 minutos y mayores a 60 minutos en un curso.

Este gráfico da al docente una “visión global” de los tipos de sesiones que se realizan en el curso a lo largo de una semana de estudio seleccionada.

6.5.5.4. Seguimiento de Tareas (Vista Docente)

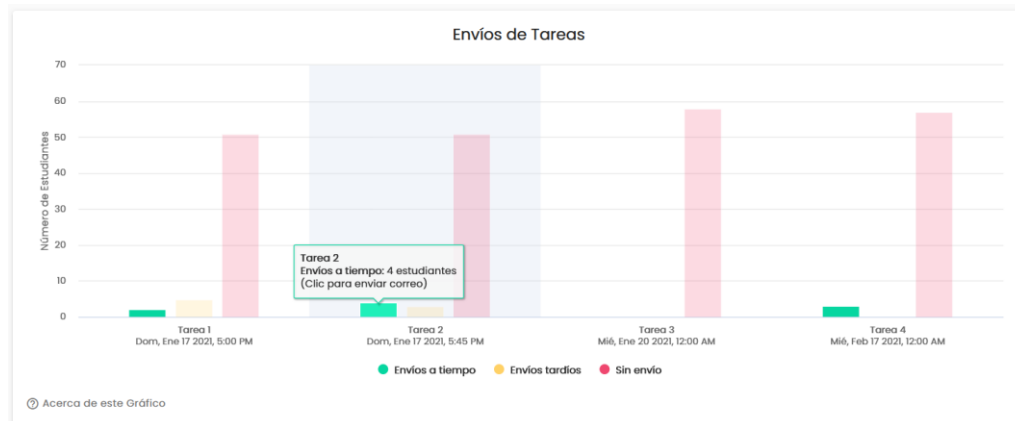
Esta sección contiene indicadores relacionados a la entrega de tareas y acceso a recursos con el objetivo de evidenciar la participación de los estudiantes con el material didáctico del curso. Los datos presentados en esta sección varían dependiendo de la semana de estudio seleccionada.

6.5.5.4.1. Envíos de Tareas

La primera visualización muestra un gráfico de barras con la distribución de la cantidad de estudiantes, respecto al estado de entrega de una tarea (Figura 74).

Figura 74

Histograma para Representar la Cantidad de Estudiantes que han enviado, que no han enviado o han enviado tarde una Tarea



Nota. Tomada de FlipMyLearning (Seguimiento de Tareas de la vista del Docente).

Este gráfico responde la siguiente pregunta:

P_SA_1. Se desea conocer la cantidad de estudiantes que han enviado a tiempo, tarde o que no han enviado actividades de tipo **assign** en cada una de las diferentes semanas planificadas a lo largo del curso. Además, permitir la posibilidad de enviar un email a los estudiantes que han enviado, no han enviado o han enviado tarde una actividad de tipo **assign**.

En el eje *x* del gráfico se encuentran las tareas de las secciones asignadas a la semana junto con la fecha y hora de entrega. En el eje *y* se encuentra la distribución del número de estudiantes según el estado de entrega. Al dar clic en alguna de las distribuciones de estudiantes se mostrará un modal con título **Enviar Correo** en su cabecera, el cuerpo del modal se asemeja a la estructura de redacción de un correo electrónico cualquiera, la sección *Para* contiene las direcciones de correo electrónico de los estudiantes en dicha distribución, el asunto del correo electrónico se completa de manera automática con el nombre del curso y el nombre de la actividad relacionada (este campo no es editable). El texto del correo electrónico puede ser cualquier descripción que el docente crea conveniente para sus estudiantes. Finalmente, el pie del modal contiene las opciones de enviar el correo electrónico o cancelar dicha acción (Figura 75).



Figura 75

Estructura del Modal de Envío de Correo Electrónico a Estudiantes

Nota. Tomada de FlipMyLearning (envío de correos)

Cuando el docente selecciona la opción de enviar, el correo electrónico empezará a enviarse a todos y cada uno de los estudiantes en la lista (Figura 76).

Figura 76

Envío de Correo Electrónico

Nota. Tomada de FlipMyLearning (envío de correos)

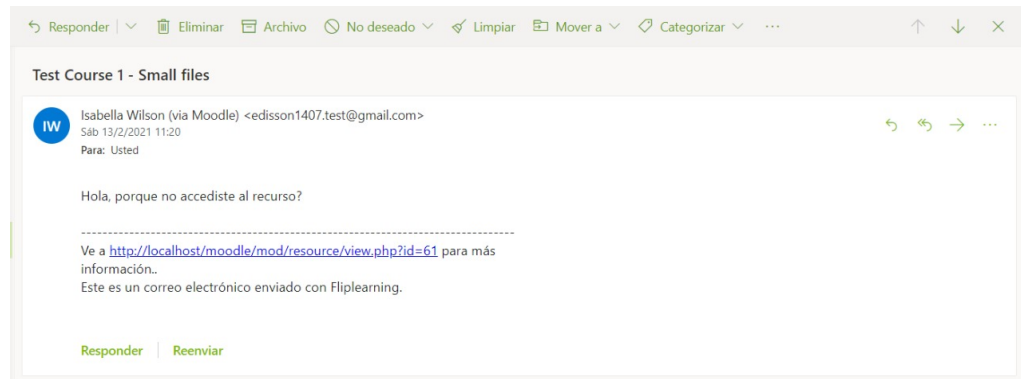
Los estudiantes recibirán un correo electrónico como se muestra en la Figura 77. Este correo se muestra con el asunto previamente explicado. Indica el texto que el docente envió, un enlace directo



a la actividad/recurso de Moodle para que el estudiante sepa con exactitud el porqué del correo electrónico y un texto indicando que el correo electrónico ha sido generado usando FlipMyLearning.

Figura 77

Correo Electrónico Generado por FlipMyLearning



Nota. Tomada de FlipMyLearning (envío de correos)

6.5.5.4.2. Acceso a los Contenidos del Curso

La segunda visualización muestra un gráfico de barras horizontal con la cantidad de estudiantes que han accedido o no a las diferentes actividades y recursos que el docente ha puesto a disposición en el curso se ha hecho uso de un gráfico de barras horizontal. En la parte superior del gráfico se tienen los distintos tipos de recursos y actividades disponibles en Moodle con la posibilidad de que el docente seleccione el tipo de recurso o actividad de cual se desea conocer el resumen (Figura 78).



Figura 78

Gráfico de barras para Evidenciar el Acceso a Recursos y Actividades de los estudiantes



Nota. Tomada de FlipMyLearning (Seguimiento de Tareas de la vista del Docente)

Esta visualización responde a la siguiente pregunta:

P_SA_2. Se desea conocer la cantidad de estudiantes que han accedido y no han accedido a los recursos y actividades (que el docente seleccione) durante las diferentes semanas de duración del curso. Además, permitir la posibilidad de enviar un email a los estudiantes que han accedido o no a dichos recursos y actividades.

En el eje x del gráfico se encuentran la cantidad de estudiantes matriculados en el curso. En el eje y del gráfico se encuentran los recursos de las secciones asignadas a la semana. Además, este gráfico permite enviar un correo electrónico a los estudiantes que han accedido al recurso o bien a aquellos que no han accedido al dar clic sobre el gráfico.

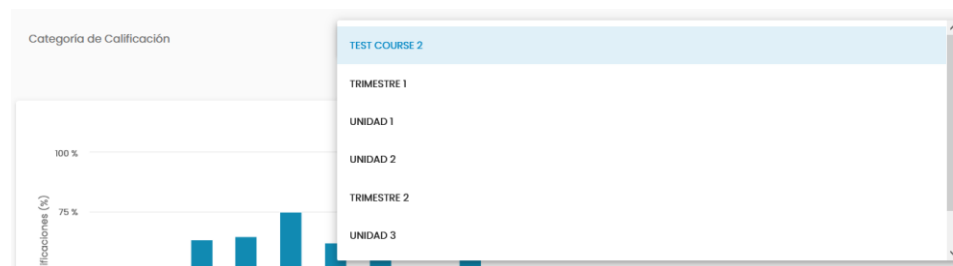
6.5.5.5. Seguimiento de calificaciones (Vista Docente)

Esta sección contiene indicadores relacionados a los promedios de calificaciones en las actividades evaluables. Las diferentes unidades didácticas (Categorías de Calificación) creadas por el docente se muestran en un selector denominado *Categoría de Calificación* (Figura 79), el cual carga

las categorías definidas en el *Libro de Calificaciones* del curso. Este selector permitirá cambiar entre las diferentes unidades definidas y mostrar las actividades evaluables en cada una. Es importante que el docente defina una planificación de las Categorías de Calificación de su curso, además de la ponderación de cada una de las actividades evaluables, de manera que el curso tenga establecida cuál es la nota máxima que un estudiante puede alcanzar.

Figura 79

Selector de Unidades Definidas para Actividades Evaluables



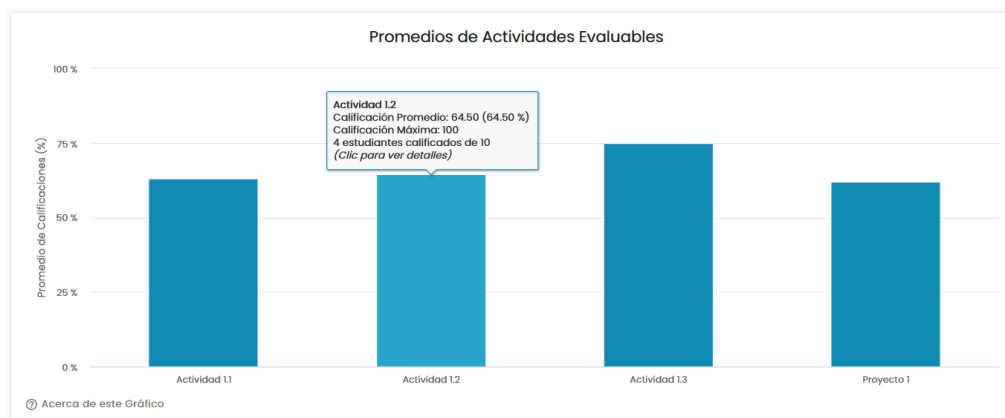
Nota. Tomada de FlipMyLearning (Seguimiento de calificaciones de la vista del Docente).

6.5.5.5.1. *Promedios de Actividades Evaluables*

La primera visualización muestra el promedio (en porcentaje) de calificaciones de los estudiantes en cada una de las actividades evaluables del curso (Figura 80) en un gráfico de barras vertical.

Figura 80

Promedio de Actividades Evaluables



Nota. Tomada de FlipMyLearning (Seguimiento de calificaciones de la vista del Docente).



Esta visualización responde la siguiente pregunta

P_SE_1. Se desea conocer el promedio general de calificaciones de cada una de las actividades evaluables del curso. Al seleccionar una actividad específica mostrar la distribución de calificaciones con el promedio, la nota máxima y la nota mínima para dicha actividad.

En el eje x del gráfico se encuentran las distintas actividades evaluables del curso. En el eje y se encuentra el promedio de calificaciones expresado en porcentaje. Se ha expresado el promedio de calificaciones en base a porcentajes puesto que Moodle permite crear actividades y asignar calificaciones máximas personalizadas. Es decir, el docente puede crear una actividad y asignar la calificación máxima como cuantitativa o cualitativa. En el caso de que la calificación sea cuantitativa, Moodle da la libertad de configurar la nota máxima de acuerdo al criterio del docente. (Por ejemplo, el docente puede crear una actividad evaluable con una calificación máxima de 100 puntos y otra con una calificación máxima de 35 puntos. Para mantener la simetría del gráfico contemplando estos dos casos de calificaciones se ha hecho uso del promedio de calificaciones basado en porcentajes, de manera que las dos actividades creadas por el docente tendrán una altura máxima de 100 en el gráfico, la cual corresponde al 100% de la calificación de cada actividad). En el caso de que las calificaciones asignadas a una actividad sean de tipo cualitativo, dichas actividades no se muestran en las visualizaciones.

Al hacer clic sobre la barra correspondiente a una actividad evaluable, los datos de los dos gráficos inferiores se actualizarán para mostrar información adicional de la actividad evaluable seleccionada.

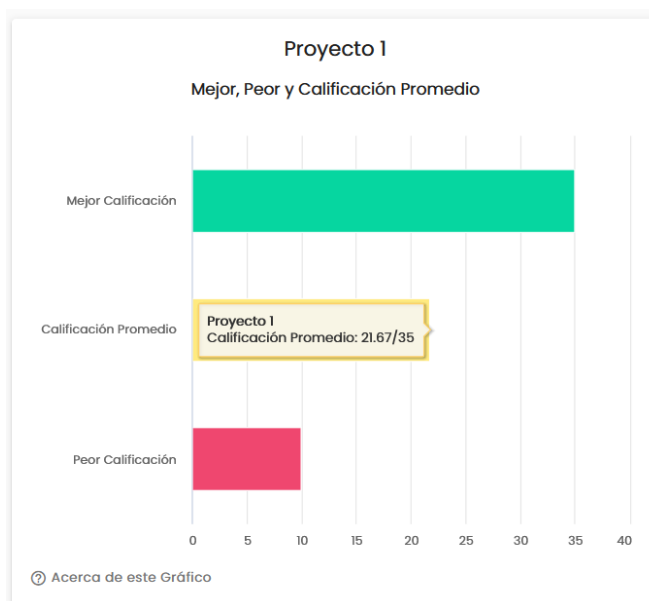
6.5.5.5.2. Mejor, Peor y Calificación Promedio

La segunda visualización muestra en un gráfico de barras horizontal el detalle de la mejor calificación, la calificación promedio y la peor calificación en la actividad seleccionada en el primer gráfico (*Promedios de Actividades Evaluables*) (Figura 80).

En el eje x del gráfico se encuentra el puntaje de la actividad para cada una de las etiquetas antes descritas, siendo la nota máxima de la actividad el máximo valor que puede alcanzar la etiqueta Mejor Calificación. Hay que tener en cuenta que en el eje x los valores varían dependiendo de la calificación máxima que el docente haya establecido previamente para la actividad. Un ejemplo se puede observar en las Figura 81 y Figura 82. En la primera, el docente ha configurado que la calificación máxima sea 35 mientras que en la segunda la calificación máxima es 100 puntos.

Figura 81

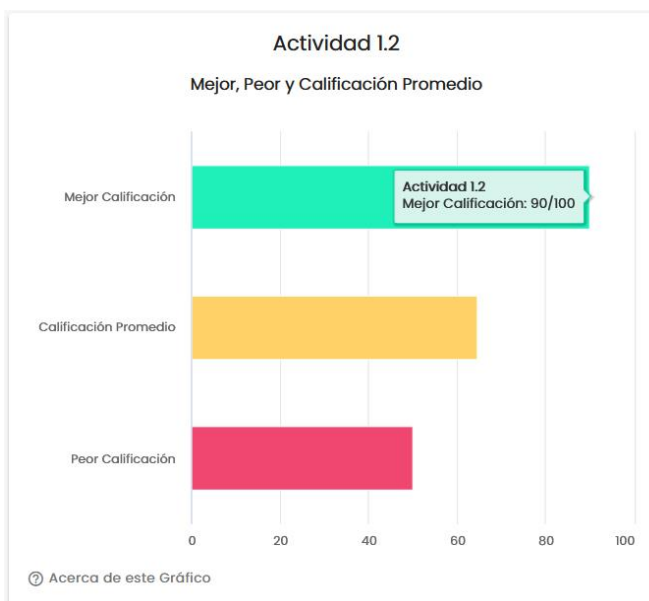
Ejemplo de actividad con ponderación máxima de 35 puntos



Nota. Tomada de FlipMyLearning (Seguimiento de calificaciones de la vista del Docente).

Figura 82

Ejemplo de Actividad con Ponderación máxima de 100 puntos



Nota. Tomada de FlipMyLearning (Seguimiento de calificaciones de la vista del Docente).



6.5.5.5.3. Distribución de Calificaciones

La tercera visualización muestra un gráfico de barras vertical con la distribución de estudiantes en diferentes rangos de calificación. Los rangos de calificación se calculan en base a porcentajes. Los rangos definidos son: menor al 50%, mayor al 50%, mayor al 60%, mayor al 70%, mayor al 80% y mayor al 90%. Estos rangos se calcularon en base a la ponderación que el docente asignó a una actividad evaluable. Un ejemplo se puede observar en la Tabla 16 y Tabla 17. Los rangos de calificación definidos en la Tabla 16 corresponden a una actividad cuya calificación máxima es 35, mientras que los rangos de la Tabla 17 corresponden a una actividad con calificación máxima de 100. La gráfica de distribución de calificaciones para la actividad con calificación máxima de 35 está en la Figura 83, mientras que la gráfica de distribución de calificaciones para la actividad con calificación máxima de 100 está en la Figura 84.

Tabla 16

Rangos de Calificaciones Calculados para la Actividad denominada Proyecto 1

Porcentaje	Rango de Calificación
Menor al 50%	Desde 0 a 17.4
Mayor al 50%	De 17.5 a 20.9
Mayor al 60%	De 21 a 24.4
Mayor al 70%	De 24.5 a 27.9
Mayor al 80%	De 28 a 31.4
Mayor al 90%	De 31.5 a 35

Tabla 17

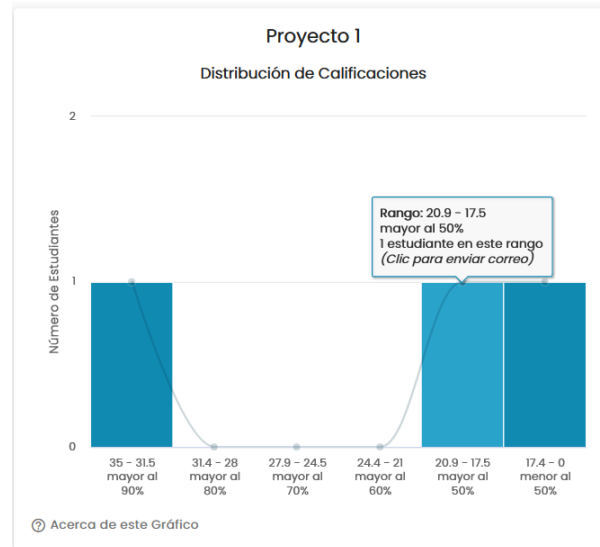
Rangos de Calificaciones Calculados para la Actividad denominada Actividad 1.2

Porcentaje	Rango de Calificación
Menor al 50%	Desde 0 a 17.4
Mayor al 50%	De 17.5 a 20.9
Mayor al 60%	De 21 a 24.4
Mayor al 70%	De 24.5 a 27.9
Mayor al 80%	De 28 a 31.4
Mayor al 90%	De 31.5 a 35



Figura 83

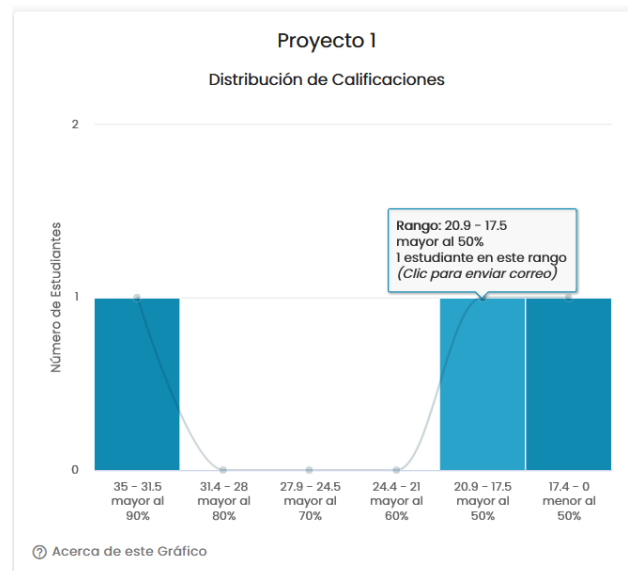
Rangos de Calificación de una Actividad con Ponderación Máxima de 35 puntos



Nota. Tomada de FlipMyLearning (Seguimiento de calificaciones de la vista del Docente).

Figura 84

Rangos de Calificación de una Actividad con Ponderación máxima de 100 puntos



Nota. Tomada de FlipMyLearning (Seguimiento de calificaciones de la vista del Docente).

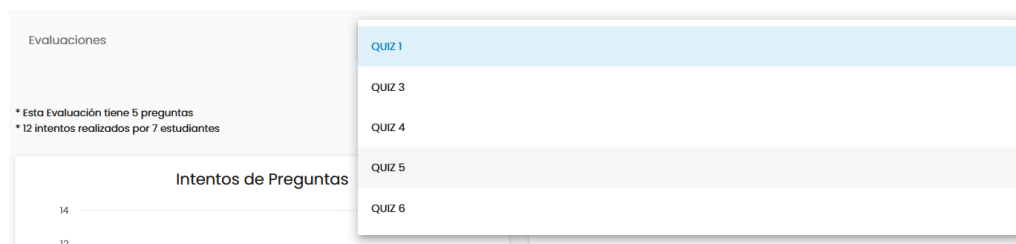
Al hacer clic sobre la barra correspondiente a un rango en las Figura 83 y Figura 84 se puede enviar un correo electrónico a los estudiantes dentro del rango de calificación.

6.5.5.6. Seguimiento de evaluaciones (Vista Docente)

Esta sección contiene indicadores relacionados al resumen de intentos en las diferentes evaluaciones del curso y análisis de preguntas de una evaluación. El análisis se lo realiza a nivel de una Evaluación individual, por lo que en la parte inferior está un selector con todas las evaluaciones de la semana seleccionada (Figura 85).

Figura 85

Selector de Evaluaciones en una Semana de Estudio

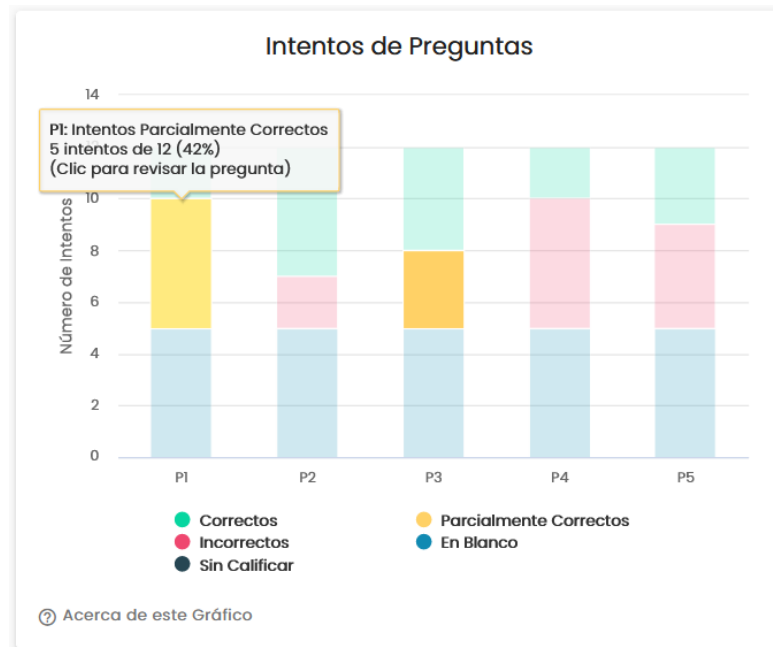


Nota. Tomada de FlipMyLearning (Seguimiento de evaluaciones de la vista del Docente).

En la parte inferior del selector de la evaluación se tiene una etiqueta que indica la cantidad de preguntas que tiene la evaluación y la cantidad de intentos de resolución de la evaluación realizados por los estudiantes. Si una evaluación tiene configurada a uno el número máximo de intentos, entonces el número de intentos será siempre igual al número de estudiantes en esta etiqueta. De lo contrario, el número de intentos será igual o mayor a la cantidad de estudiantes que realizaron la evaluación.

6.5.5.6.1. Intentos de Preguntas

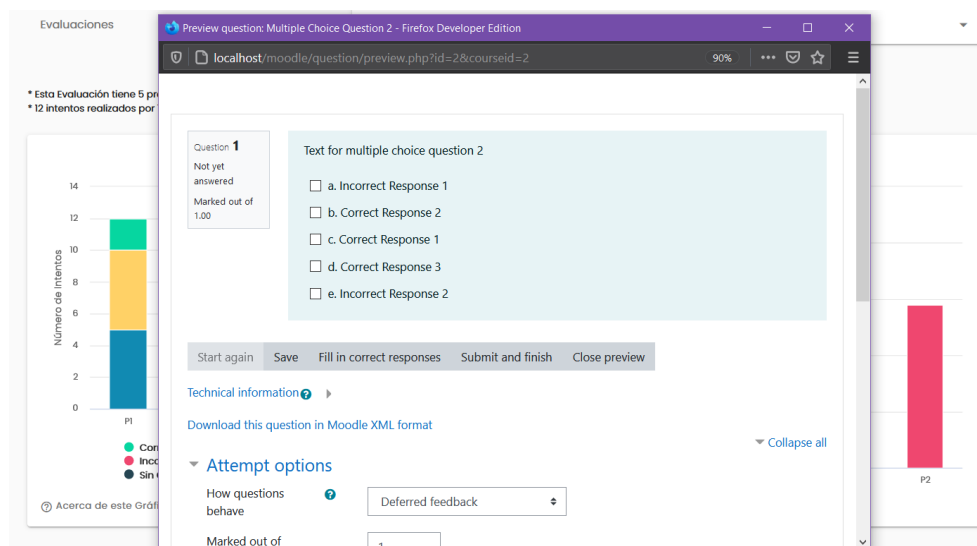
La primera visualización muestra en un gráfico de barras apilado vertical la distribución de intentos de cada una de las preguntas de una evaluación (Figura 86).

Figura 86*Histograma de Intentos de Preguntas para una evaluación*

Nota. Tomada de FlipMyLearning (Seguimiento de evaluaciones de la vista del Docente).

En el eje x del gráfico se encuentran las preguntas de la evaluación. En el eje y se encuentra la cantidad de intentos de resolución para cada una de dichas preguntas. La simetría del gráfico se verá afectada por la configuración de la evaluación (en una evaluación que tenga siempre las mismas preguntas, el gráfico presentará la misma cantidad de intentos para cada barra correspondiente a una pregunta. En una evaluación que tenga preguntas aleatorias (de un banco de preguntas), el gráfico presentará en la barra de cada pregunta la suma de los intentos de evaluaciones en los que apareció, pudiendo no ser la misma para cada pregunta de la evaluación).

Al hacer clic en alguna de las barras correspondiente a una pregunta es posible ver la pregunta de la evaluación en una ventana emergente (Figura 87).

Figura 87*Detalle de la Pregunta en la Evaluación*

Nota. Tomada de FlipMyLearning (Seguimiento de evaluaciones de la vista del Docente).

6.5.5.6.2. Preguntas más Difíciles

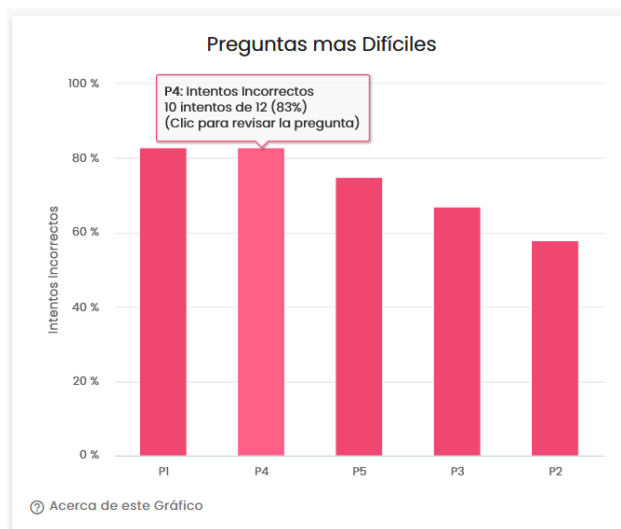
La segunda visualización muestra en un gráfico de barras vertical con las preguntas de la evaluación ordenadas por su nivel de dificultad (Figura 88). Se considera incorrecto a un intento de resolución de una pregunta con el estado de Parcialmente Correcto, Incorrecto o En Blanco, de manera que la cantidad total de intentos incorrectos de una pregunta es la suma de los intentos con los estados antes mencionados. El nivel de dificultad se representa en porcentaje calculado en base a la cantidad total de intentos.

En el eje x del gráfico se encuentran las preguntas de la evaluación identificadas por el nombre. En el eje y se encuentran el porcentaje de intentos incorrectos del total de intentos de la pregunta. Este eje permite identificar cuáles han sido las preguntas que han representado mayor dificultad para los estudiantes que rindieron la evaluación.

Al hacer clic en alguna de las barras correspondiente a una pregunta es posible ver la pregunta de la evaluación en una ventana emergente (Figura 87).

Figura 88

Nivel de Dificultad de cada una de las Preguntas de una Evaluación



Nota. Tomada de FlipMyLearning (Seguimiento de evaluaciones de la vista del Docente).

6.5.5.7. Deserción (Vista Docente)

Esta sección contiene indicadores relacionados a la predicción de abandono de estudiantes de un curso. La información se muestra en base a grupos de estudiantes calculados por un algoritmo de Clustering que analiza el comportamiento de cada estudiante en base al tiempo invertido, la cantidad de sesiones del estudiante, la cantidad de días activo y las interacciones que ha realizado con cada recurso y con los demás estudiantes del curso ([Detección de las variables de predicción de abandono en Moodle](#)). El algoritmo coloca en el mismo grupo a estudiantes con similar comportamiento, de manera que se puede identificar a los estudiantes más y menos comprometidos con el curso. Los datos presentados en esta sección varían dependiendo del grupo seleccionado en el selector que contiene los grupos identificados en el curso (Figura 89).

Figura 89

Selector de Grupos de Estudiantes Identificados

Grupo de Estudiantes

Grupo 2

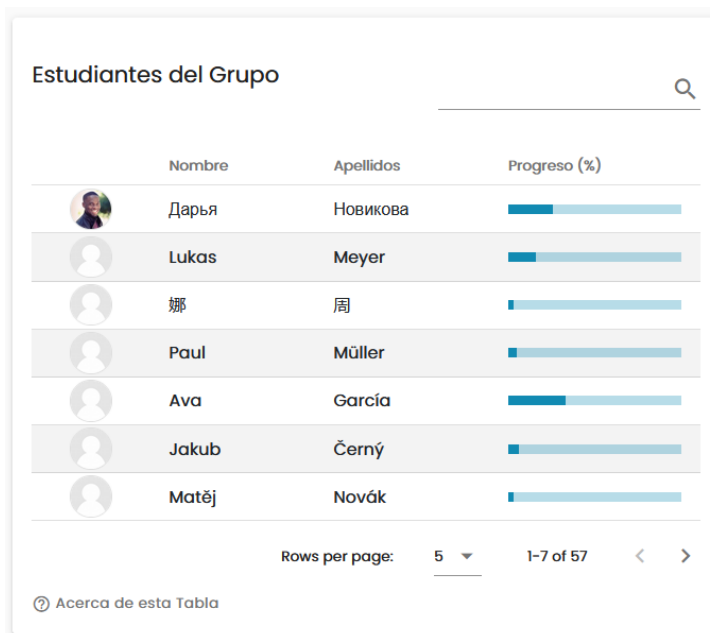
Nota. Tomada de FlipMyLearning (Deserción de la vista del Docente).








6.5.5.7.1. *Estudiantes del Grupo*

La primera de las visualizaciones se denomina *Estudiantes del Grupo* y presenta una tabla interactiva con datos de cada uno de los estudiantes pertenecientes al grupo (clúster generado por el algoritmo). Esta tabla muestra la foto del perfil del estudiante (en caso de estar disponible), nombres y apellidos de los estudiantes y el progreso en el curso. El progreso calculado sigue la misma lógica que la explicada en la sección de [Sesiones de Estudio](#). Además, la tabla posee acciones como el ordenamiento de los estudiantes por nombres, apellidos o progreso. Búsqueda de un estudiante y paginación de registros para una mejor visualización (Figura 90). Además, esta tabla posee el evento de clic, el cual permite actualizar todas las gráficas inferiores con los datos del estudiante seleccionado.

Figura 90

Tabla Interactiva de Estudiantes del Grupo



	Nombre	Apellidos	Progreso (%)
	Дарья	Новикова	<div><div></div></div>
	Lukas	Meyer	<div><div></div></div>
	娜	周	<div><div></div></div>
	Paul	Müller	<div><div></div></div>
	Ava	García	<div><div></div></div>
	Jakub	Černý	<div><div></div></div>
	Matěj	Novák	<div><div></div></div>

Rows per page: 5 1-7 of 57 < >

[? Acerca de esta Tabla](#)

Nota. Tomada de FlipMyLearning (Deserción de la vista del Docente).

6.5.5.7.2. Perfil del Estudiante

La segunda de las visualizaciones muestra información sobre el estudiante seleccionado en la tabla de *Estudiantes del Grupo* (Figura 90). Esta visualización muestra la foto del perfil del estudiante (en caso de estar disponible), nombres y apellidos, tiempo invertido, número de sesiones, calificación en el curso, nombre de usuario en el sistema, correo electrónico, la última fecha acceso al curso y el progreso del estudiante en el curso (basado en los recursos finalizados) (Figura 91).

Figura 91

Información del Perfil del Estudiante



Nota. Tomada de FlipMyLearning (Deserción de la vista del Docente).

El progreso calculado sigue la misma lógica que la explicada en la sección de [Sesiones de Estudio](#). Además, se indica el número de módulos finalizados del total de módulos disponibles en el curso. El tiempo invertido del estudiante refleja la cantidad total de tiempo que ha permanecido en el curso. La cantidad de sesiones muestran cuántas sesiones el estudiante ha realizado en el curso (como ya se explicó, las sesiones se identificaron en función de los accesos al curso y el tiempo de inactividad entre interacciones). La calificación se obtiene mediante la siguiente fórmula:

$$\text{calificación} = \frac{\text{suma de calificaciones obtenidas en actividades}}{\text{suma de ponderaciones máximas en actividades}}$$

Si el docente no ha configurado el Libro de Calificaciones de Moodle, puede darse un caso como el mostrado en la Figura 91. Sin embargo, si el Libro de Calificaciones es previamente configurado, la calificación mostrada tendrá resultados más relevantes.

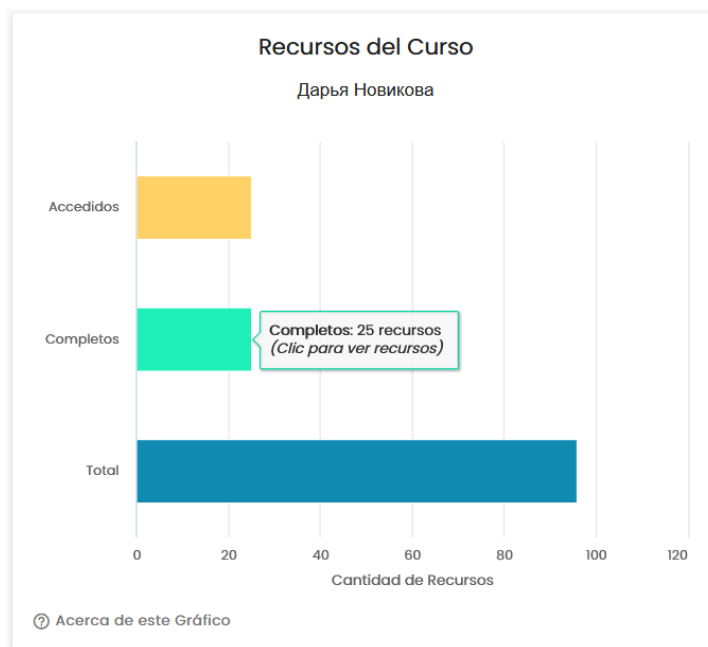
Además, la visualización posee acciones adicionales como ver el perfil del estudiante, enviar un correo electrónico a dicho estudiante o al grupo al que pertenece dicho estudiante.

6.5.5.7.3. Recursos del Curso

La tercera visualización muestra en un gráfico de barras horizontal la cantidad de recursos y actividades a las que el estudiante ha accedido o completado. En el eje x del gráfico se encuentran la cantidad de recursos y actividades que el docente ha puesto a disposición de los estudiantes en el curso. En el eje y se encuentran las etiquetas de Accedidos, Completos y Total (Figura 92).

Figura 92

Recursos divididos por etiquetas de Accedidos, Completos y Total.



Nota. Tomada de FlipMyLearning (Deserción de la vista del Docente).



Para determinar la cantidad de recursos y actividades completas se hace uso de la configuración de Moodle denominada Finalización de Actividad (*Activity Completion en inglés*), como se explica en la sección de [Sesiones de Estudio](#). Por ejemplo, si el docente creó una actividad de tipo Foro (*forum*) y configuró que el criterio de finalización sea que el estudiante realice una publicación y este no la hizo, entonces el gráfico mostrará que el estudiante accedió a dicha actividad, sin embargo, no la completó. En caso de que el docente no realice la configuración de completitud para las actividades del curso, la cantidad de actividades accedidas y completas siempre será la misma. FlipMyLearning detecta que el recurso no tiene configurado un criterio de finalización y asume que el estudiante completó la actividad si la vio al menos una vez.

Al hacer clic sobre alguna barra es posible ver los recursos y actividades disponibles en el curso (en una ventana modal) junto con la cantidad de interacciones del estudiante con cada recurso y una etiqueta de no accedido, accedido o completado (Figura 93). Además, cada uno de estos módulos contiene un enlace directo al recurso o actividad de Moodle.

Figura 93

Recursos del Curso.

Recursos del Curso		
Tema 0		
Forum (0 interacciones)	No accedido	
Social forum (1 interacción)	Accedido	Completado
Tema 1		
Small files (1 interacción)	Accedido	Completado
Assignment 1 (4 interacciones)	Accedido	Completado
Tarea 1 (12 interacciones)	Accedido	Completado
Tarea 2 (2 interacciones)	Accedido	Completado
Tarea 3 (14 interacciones)	Accedido	Completado
Page 1 (2 interacciones)	Accedido	Completado
Page 2 (1 interacción)	Accedido	Completado
Page 3 (1 interacción)	Accedido	Completado

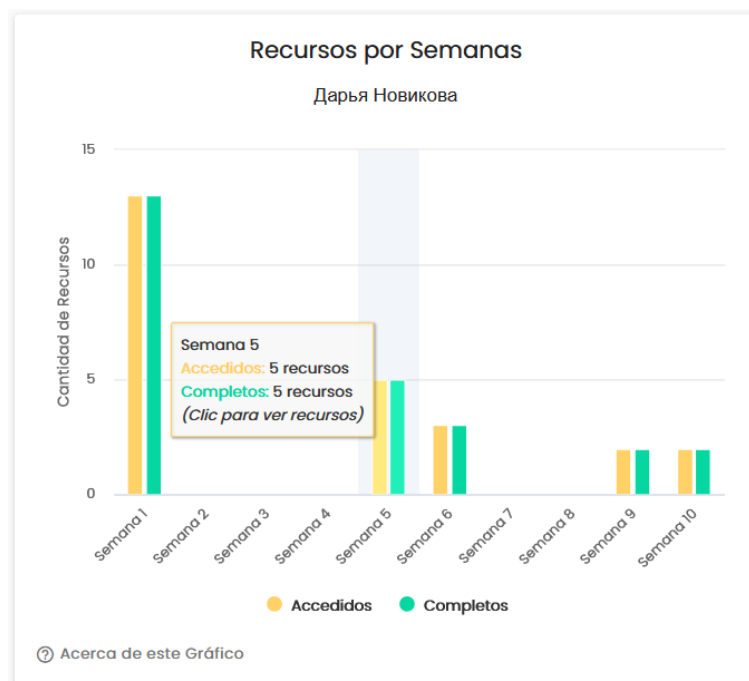
Nota. Tomada de FlipMyLearning (Deserción de la vista del Docente).

6.5.5.7.4. Recursos por Semana

La cuarta visualización muestra en un gráfico de barras vertical la cantidad de recursos y actividades que el estudiante ha accedido a lo largo de las diferentes semanas configuradas previamente mediante un gráfico de barras o histograma. En el eje x del gráfico se encuentran las diferentes semanas de estudio configuradas. En el eje y se encuentra la cantidad de recursos y actividades asignadas a cada semana (Figura 94).

Figura 94

Recursos de las diferentes Semanas de Estudio



Nota. Tomada de FlipMyLearning (Deserción de la vista del Docente).

Al hacer clic sobre alguna barra es posible ver los recursos y actividades disponibles en el curso (en una ventana emergente) junto con la cantidad de interacciones del estudiante con cada recurso y una etiqueta de no accedido, accedido o completado.

6.5.5.7.5. Sesiones y Tiempo Invertido

La quinta visualización muestra en un gráfico de líneas cómo ha evolucionado la cantidad de sesiones de estudio y el tiempo invertido desde que se registró la primera sesión del estudiante (Figura 95).

Figura 95

Evolución de las Sesiones y Tiempo Invertido de un Estudiante



Nota. Tomada de FlipMyLearning (Deserción de la vista del Docente).

En el eje x del gráfico se muestra una línea temporal que muestra los días que el estudiante ha realizado sesiones de estudio en la plataforma. Esta línea temporal muestra fechas desde el inicio del curso hasta la última sesión registrada de dicho estudiante. En el eje y se encuentra el número de sesiones del estudiante a lado izquierdo y el tiempo invertido en horas a lado derecho. Entre dichos ejes se dibujan la cantidad de sesiones y el tiempo invertido del estudiante analizado.

Esta visualización permite hacer un acercamiento sobre una región seleccionada. Este acercamiento ayuda a evidenciar de manera clara dicha evolución en diferentes rangos de fechas (Figura 96).

Figura 96

Acercamiento de Evolución de las Sesiones y Tiempo Invertido de un Estudiante en un rango de Tiempo



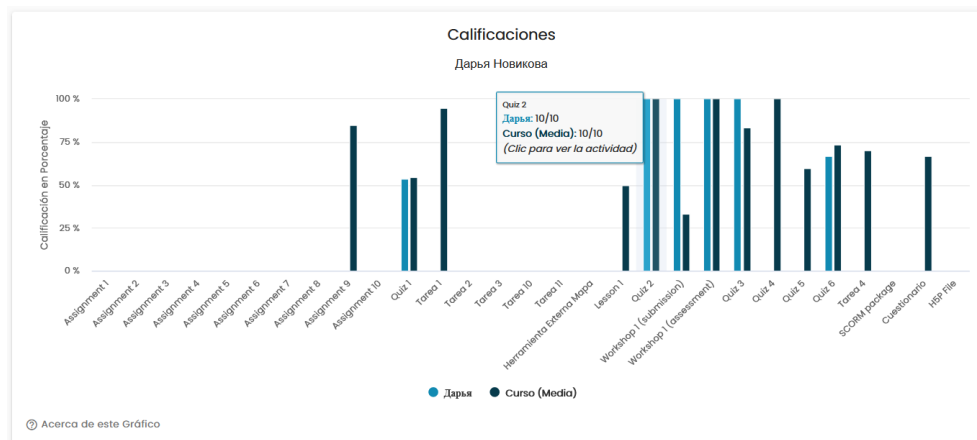
Nota. Tomada de FlipMyLearning (Deserción de la vista del Docente).

6.5.5.7.6. Calificaciones

La última visualización muestra en un gráfico de barras vertical una comparación de las calificaciones del estudiante con los promedios de calificaciones (media) de sus compañeros en las distintas actividades evaluables del curso (Figura 97). En el eje *x* del gráfico se muestran las diferentes actividades evaluables puestas a disposición del estudiante por el docente. En el eje *y* se encuentra la calificación del estudiante y la media de calificaciones de sus compañeros. Al hacer clic sobre una barra es posible dirigirse a la actividad.

Figura 97

Comparación de Calificaciones en las diferentes Actividades Evaluables



Nota. Tomada de FlipMyLearning (Deserción de la vista del Docente).

6.5.5.8. Indicadores Generales (Vista Estudiante)

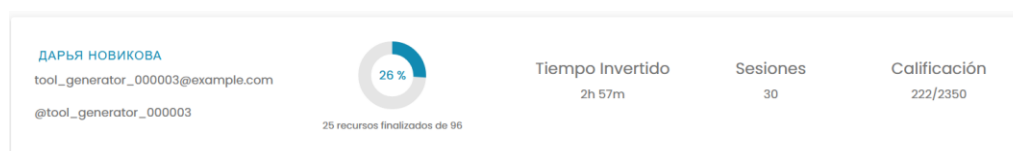
Esta sección contiene indicadores relacionados a la información, progreso, indicadores generales, recursos y actividades del curso, sesiones a lo largo de la semana, tiempo invertido y calificaciones del estudiante. Esta sección contiene cinco visualizaciones.

6.5.5.8.1. Perfil del Estudiante

La primera visualización muestra nombres y apellidos del estudiante, nombre de usuario en el sistema, correo electrónico, progreso del estudiante (El progreso se calculó siguiendo la misma lógica que la explicada en la sección de [Sesiones de Estudio](#)) (Figura 98), el tiempo invertido, el número de sesiones y la calificación en el curso (la calificación se obtiene mediante la fórmula explicada en la sección de [Deserción](#)).

Figura 98

Información del Perfil del Estudiante



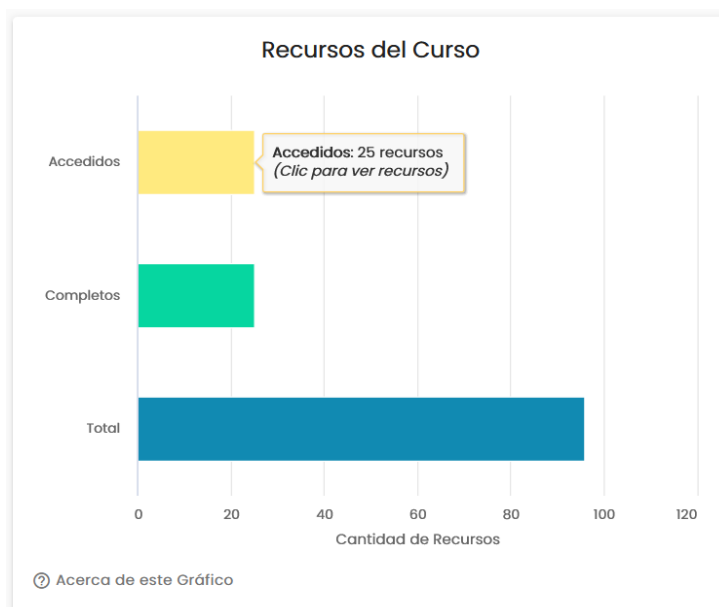
Nota. Tomada de FlipMyLearning (Indicadores Generales de la vista del Estudiante).

6.5.5.8.2. Recursos del Curso

La segunda visualización muestra en un gráfico de barras horizontal la cantidad de recursos y actividades a las que el estudiante ha accedido o completado. En el eje x del gráfico se encuentran la cantidad de recursos y actividades que el docente ha puesto a disposición de los estudiantes en el curso. En el eje y se encuentran las etiquetas de Accedidos, Completos y Total (Figura 99).

Figura 99

Recursos divididos por etiquetas de Accedidos, Completos y Total



Nota. Tomada de FlipMyLearning (Indicadores Generales de la vista del Estudiante).

Para determinar la cantidad de recursos y actividades completas se hace uso de la configuración de Moodle denominada Finalización de Actividad (*Activity Completion en inglés*). Como se explica en la sección de [Sesiones de Estudio](#) y [Deserción](#).

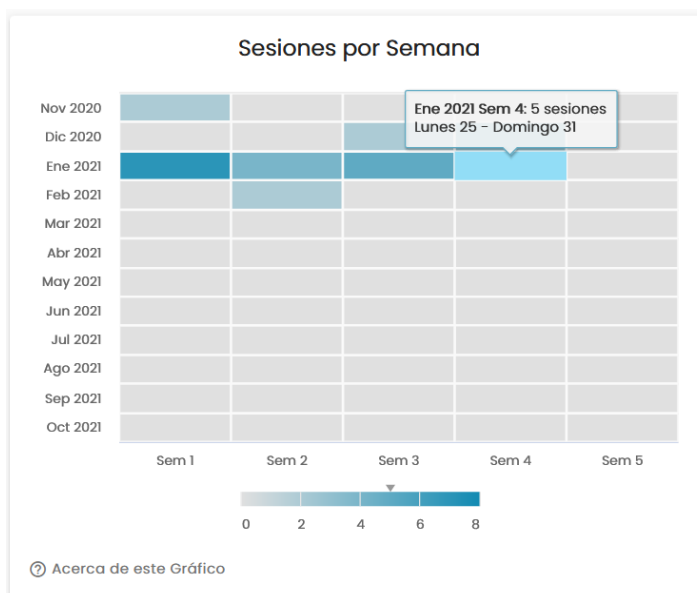
6.5.5.8.3. Sesiones por Semana

La tercera visualización muestra en un mapa de calor (*heat map en inglés*) las sesiones de estudio del estudiante a lo largo de los diferentes meses del año desde que inició el curso (Figura 100). En el eje x del gráfico se encuentran las semanas de cada mes. En el eje y del gráfico se encuentran los diferentes meses del año partiendo del mes de creación del curso. Para mantener la

simetría del gráfico se ha colocado un total de cinco semanas para cada mes, sin embargo, no todos los meses tiene tal cantidad de semanas. Dichos meses, solo sumarán sesiones hasta la semana cuatro.

Figura 100

Mapa de calor de Sesiones por Semana



Nota. Mapa de calor para evidenciar las sesiones de estudio por semanas del mes. Tomada de FlipMyLearning (Indicadores Generales de la vista del Estudiante)

6.5.5.8.4. Sesiones y Tiempo Invertido

La quinta visualización muestra en un gráfico de líneas cómo ha evolucionado la cantidad de sesiones de estudio y el tiempo invertido desde que se registró la primera sesión del estudiante (Figura 101).

Figura 101*Evolución de las Sesiones y Tiempo Invertido de un Estudiante*

Nota. Tomada de FlipMyLearning (Indicadores Generales de la vista del Estudiante).

En el eje x del gráfico se muestra una línea temporal que muestra los días que el estudiante ha realizado sesiones de estudio en la plataforma. Esta línea temporal muestra fechas desde el inicio del curso hasta la última sesión registrada de dicho estudiante. En el eje y se encuentra el número de sesiones del estudiante a lado izquierdo y el tiempo invertido en horas a lado derecho. Entre dichos ejes se dibujan la cantidad de sesiones y el tiempo invertido del estudiante analizado. Esta visualización permite hacer un acercamiento sobre una región seleccionada. Este acercamiento ayuda a evidenciar de manera clara dicha evolución en diferentes rangos de fechas.

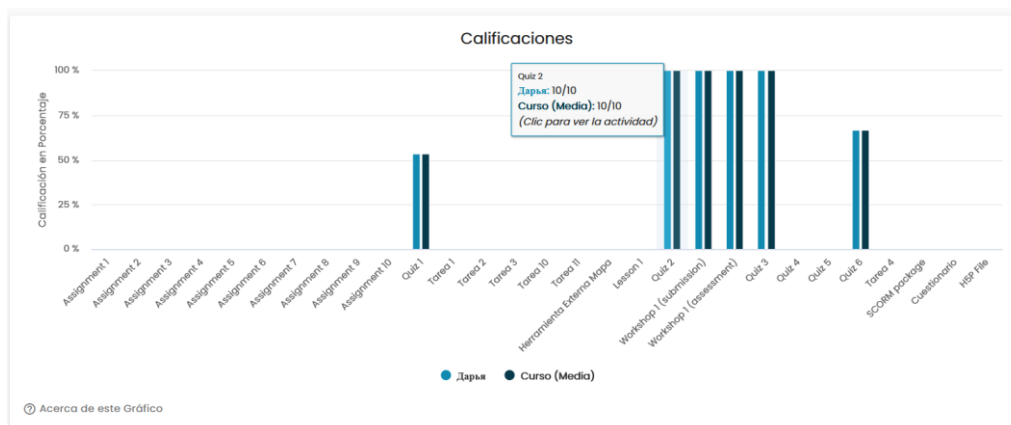
6.5.5.8.5. Calificaciones

La última visualización muestra en un gráfico de barras vertical una comparación de las calificaciones del estudiante con los promedios de calificaciones (media) de sus compañeros en las distintas actividades evaluables del curso (Figura 102). En el eje x del gráfico se muestran las diferentes actividades evaluables puestas a disposición del estudiante por el docente. En el eje y se encuentra la calificación del estudiante y la media de calificaciones de sus compañeros. Al hacer clic sobre una barra es posible dirigirse a la actividad.



Figura 102

Comparación de Calificaciones en las diferentes Actividades Evaluables



Nota. Tomada de FlipMyLearning (Indicadores Generales de la vista del Estudiante).

6.5.5.9. Sesiones de Estudio (Vista Estudiante)

Esta sección contiene visualizaciones con indicadores relacionados a la actividad del estudiante en el curso medida en tiempo invertido en el curso, sesiones realizadas e interacción con los recursos del curso. Los datos de esta sección dependen de la semana de estudio seleccionada.

6.5.5.9.1. Tiempo Invertido

La primera visualización muestra el tiempo invertido por parte del estudiante en la semana seleccionada en comparación del tiempo planificado por parte del docente en forma de un gráfico de barras horizontal (Figura 103).

Figura 103

Gráfico de barras para evidenciar el Tiempo promedio invertido del Estudiante a Comparación del Tiempo promedio planificado por el Docente a lo largo de una semana de estudio



Nota. Tomada de FlipMyLearning (Sesiones de Estudio de la vista del Estudiante).

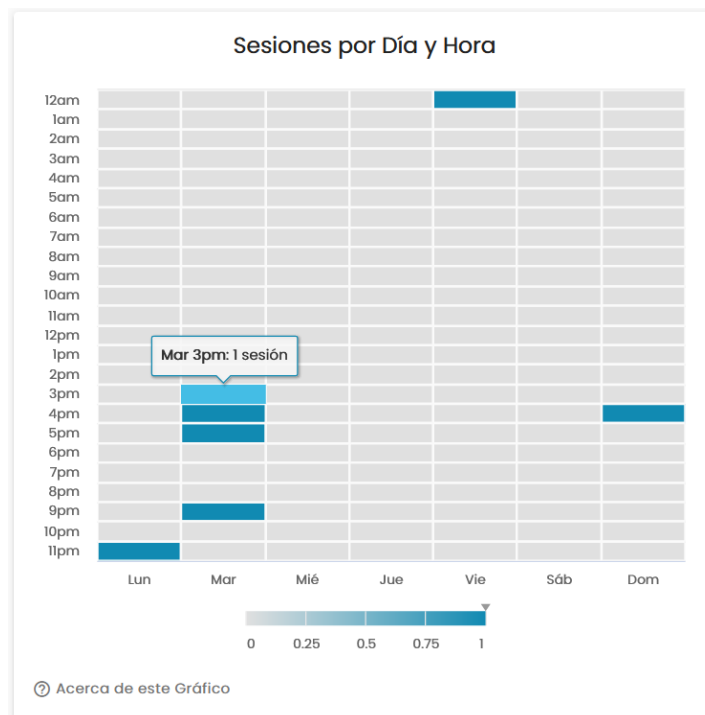
En el eje x del gráfico se encuentra el número de horas que el docente ha planificado para una semana específica y en el eje y se encuentran las etiquetas de tiempo promedio invertido y tiempo promedio que se debería invertir.

6.5.5.9.2. Sesiones por Día y Hora

La segunda visualización muestra un gráfico con las sesiones de estudio del estudiante por día y hora de la semana seleccionada en un mapa de calor (Figura 104).

Figura 104

Mapa de calor de Sesiones por Día y Hora

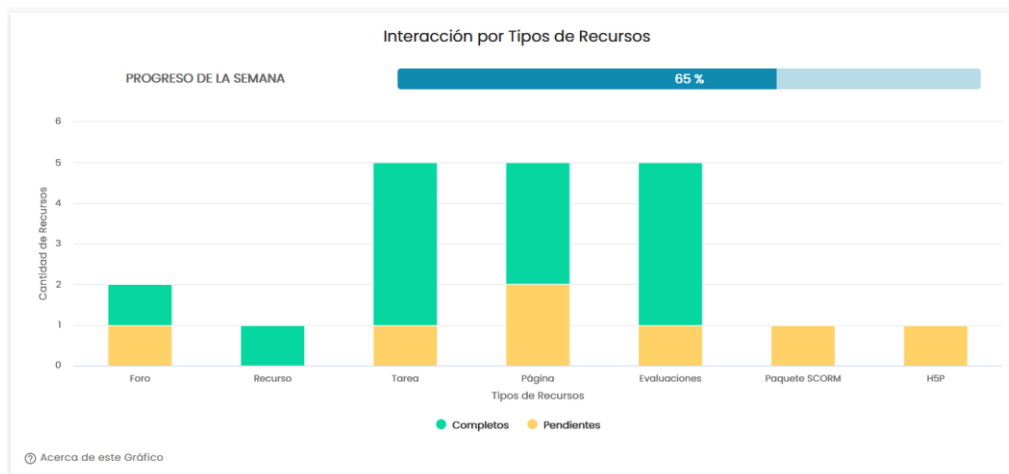


Nota. Tomada de FlipMyLearning (Sesiones de Estudio de la vista del Estudiante).

En el eje x del gráfico se encuentran los diferentes días de la semana y en el eje y se encuentran las horas del día empezando por las 12am y terminando a las 11pm o 23 horas.

6.5.5.9.3. Interacciones por Tipo de Recursos

La última de las visualizaciones muestra en un gráfico de barras apilado vertical los recursos o actividades pendientes o completas en cada una de las semanas de estudio. En el eje x del gráfico se encuentran los diferentes tipos de recursos o actividades disponibles en el curso. En el eje y se encuentran la cantidad de recursos o actividades (Figura 105).

Figura 105*Histograma Apilado de Interacción por Tipos de Recursos*

Nota. Tomada de FlipMyLearning (Sesiones de Estudio de la vista del Estudiante).

Al hacer clic en la barra seleccionada es posible visualizar los recursos o actividades de interés como se muestra en la Figura 94. Además, el estudiante puede observar una barra de progreso denominada Progreso de la Semana, el cual es calculado en base a las actividades y recursos completados en las diferentes semanas de estudio configuradas por el docente.

6.5.6 Funciones adicionales

FlipMyLearning trabaja con la característica de Grupos de Moodle. El uso de grupos es una gran estrategia de colaboración y diferenciación. En Moodle, se puede organizar a los estudiantes en distintos grupos. Al asignar estudiantes a grupos, el docente puede permitir trabajar en colaboración en un entorno cerrado (los miembros del grupo ven solo el trabajo de su grupo) o abierto (los miembros del grupo ven el trabajo de otros grupos) (Moodle, 2020). El uso de grupos también facilita la gestión del trabajo del docente de organizar, revisar y evaluar a los estudiantes. FlipMyLearning permite filtrar los datos de las visualizaciones mediante los grupos definidos en el curso (en caso de que el docente haya creado grupos de estudiantes). Si el docente no ha creado grupos de estudiantes, los resultados de las visualizaciones serán de todos los estudiantes que pertenecen al curso. Esta característica está presente en todas las secciones de FlipMyLearning a excepción de la sección

[Configurar semanas](#). Este filtro puede ser usado por el docente en la parte superior de cada una de las secciones como se muestra en la Figura 106. Se ha hecho uso de una lista desplegable (recurso que permite al usuario seleccionar una opción) que contendrá los diferentes grupos (en caso de existir). La opción por defecto será TODOS LOS ESTUDIANTES, que hace referencia a que los datos de las visualizaciones toman en cuenta a todos los estudiantes del curso.

Figura 106

Usar Filtro de Grupos del Curso



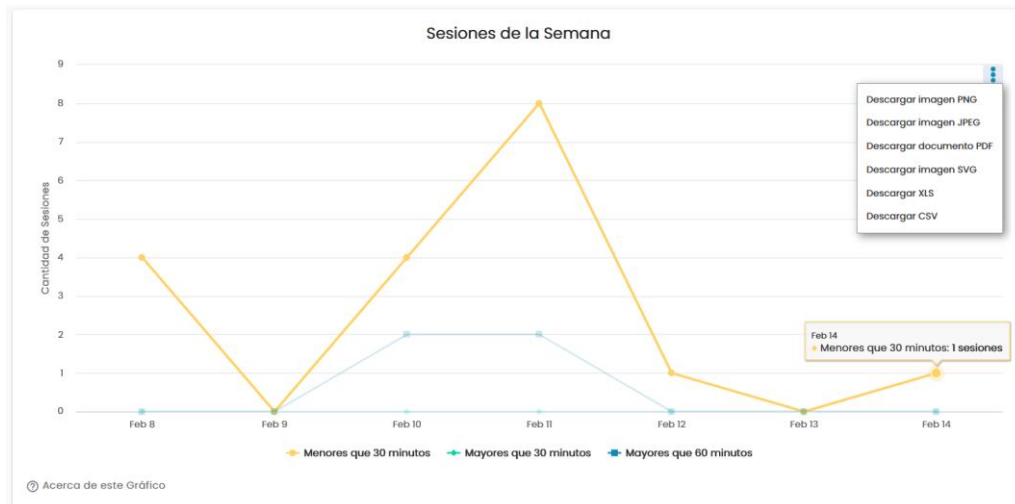
Nota. Tomada de FlipMyLearning (Sesiones de estudio).

Gracias a la librería Highcharts es posible obtener funciones adicionales para cada uno de los gráficos antes mencionados con excepción de la tabla interactiva denominada Progreso del Curso de la Figura 75. Entre las funciones integradas se incorpora un módulo de exportación. Este módulo permite a sus usuarios descargar el gráfico como PDF, PNG, JPEG o imágenes vectoriales SVG. Incluso exportar los datos del gráfico a formatos de tabla CSV, XLS. El formato de datos del archivo dependerá del gráfico seleccionado. Todas estas funciones están disponibles en cada uno de los gráficos en un menú personalizado como se muestra en la Figura 107.

Adicional a todo lo descrito anteriormente, las visualizaciones de FlipMyLearning han sido creadas con un diseño web responsive o adaptativo. De modo que, las visualizaciones puedan ser percibidas de la mejor manera posible en distintos dispositivos (desde ordenadores de escritorio a tablets y móviles). Esto permite una mejor experiencia e interacción del usuario con la herramienta.

Figura 107

Menú de Opciones Adicionales para los Gráficos de FlipMyLearning



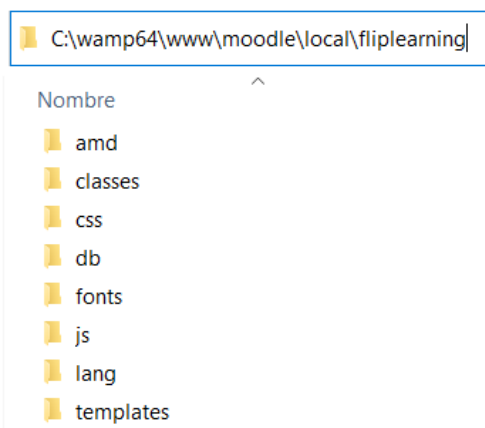
Nota. Tomada de FlipMyLearning (Sesiones de estudio).

6.5.7. Estructura de directorios

La estructura de directorios de la herramienta desarrollada sigue la organización que recomienda Moodle para el desarrollo de un plugin externo (Figura 108).

Figura 108

Estructura de directorios de FlipMyLearning



Nota. Tomada de FlipMyLearning (Estructura de directorios).

FlipMyLearning es una extensión de tipo local en la que se distinguen los siguientes directorios:

- */amd*: contiene archivos de javascript minificados.
- */clases*: contiene las clases de php con la lógica de programación de las funcionalidades del plugin.
- */css*: contiene archivos css utilizados para estilizar las vistas web del plugin.
- */db*: contiene script de definición de tablas de base de datos útiles cuando se instala el plugin.
- */fonts*: contiene los archivos de las tipografías utilizadas en el plugin.
- */js*: contiene las librerías de javascript utilizadas en el plugin.
- */lang*: contiene archivos con las definiciones de todo el texto para la traducción del plugin a diferentes idiomas.
- */templates*: contiene los archivos que renderizan las vistas web del plugin.

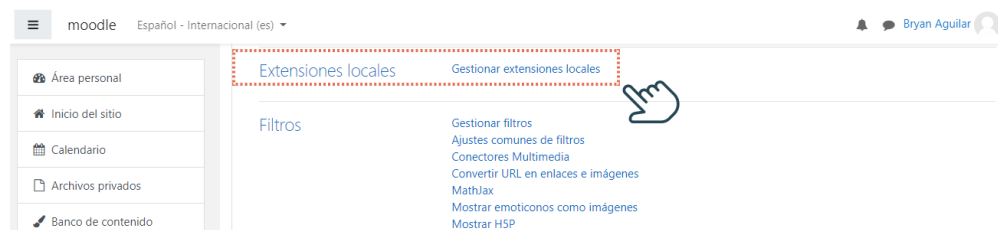
6.5.8. Proceso de desinstalación

FlipMyLearning tiene un proceso de desinstalación sencilla. Se desinstala como cualquier otra extensión/plugin de Moodle. Para ejecutar este proceso es necesario tener una cuenta de administrador del sitio y seguir los siguientes pasos:

- Iniciar sesión como administrador de Moodle. Posteriormente, en la pantalla inicial luego de iniciar sesión, ingresar a la opción de Administración del Sitio >Extensiones >Extensiones locales y abrir la opción *Gestionar extensiones locales* como se muestra en la Figura 109.

Figura 109

Extensiones locales de Moodle



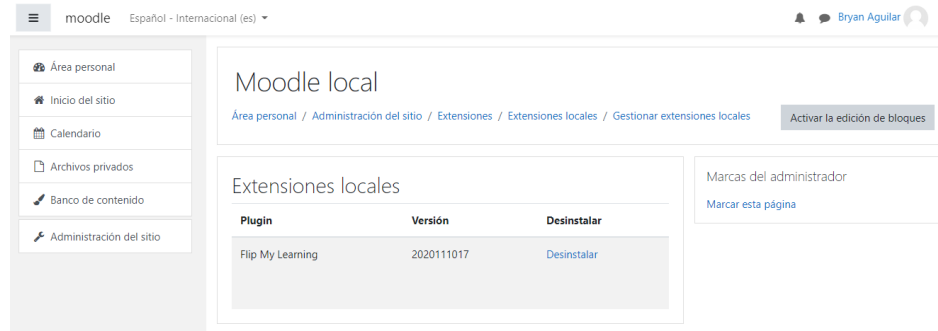
Nota. Elaboración Propia.



- A continuación, se muestra una lista con las extensiones locales instaladas en Moodle. En este caso únicamente se muestra FlipMyLearning ya que es la única extensión local disponible. Posteriormente, se selecciona la opción *Desinstalar*.

Figura 110

Lista de extensiones locales Disponibles en Moodle

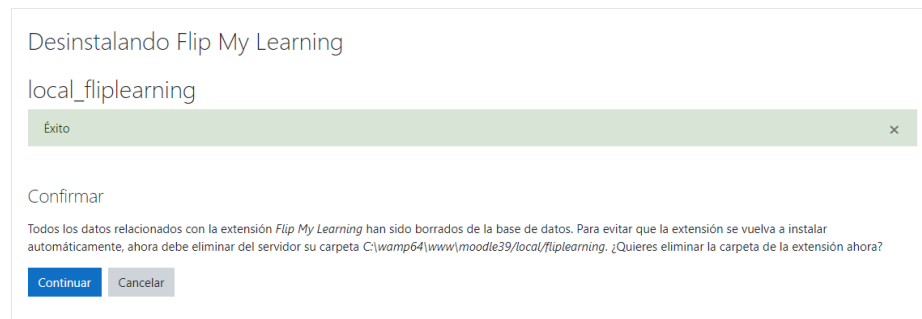


Nota. Elaboración Propia.

- Al seleccionar la opción Desinstalar Moodle muestra un mensaje de confirmación de la desinstalación. Este proceso borra el directorio de la carpeta `local` de Moodle, además, se encarga de eliminar todas las tablas de datos creadas en la etapa de instalación. Una vez finalizado este proceso Moodle indicará que el proceso de desinstalación se completó de manera correcta.

Figura 111

Proceso de Desinstalación de FlipMyLearning



Nota. Elaboración Propia.



6.6. Evaluación del Dashboard

Para la evaluación de usuarios, se ha convocado a 12 participantes. Dichos participantes forman parte de un curso virtual en Moodle, cinco con el rol de docente y siete con el rol de estudiante. Los participantes tuvieron acceso completo a FlipMyLearning, tanto para la vista del estudiante como a la vista para el docente. Se ha hecho uso de dos cuestionarios, uno para cada vista. A cada participante se le proporcionó un enlace a un formulario en línea. Dicho formulario fue realizado con la herramienta Google Forms.

Para una medida estándar de la experiencia de los participantes, además de la importancia de las métricas y visualizaciones con FlipMyLearning, se ha hecho uso del conocido Cuestionario de Experiencia de Usuario (*UEQ – User Experience Questionnaire en inglés*) (Laugwitz, Held, & Schrepp, 2008). Consiste en elementos clasificados en una escala Likert que varía de 1 (totalmente en desacuerdo) a 5 (totalmente de acuerdo), para medir factores de la experiencia del usuario (Figura #) como:

- *Atractivo*: impresión general. ¿A los usuarios les gusta o no les gusta?
- *Claridad*: ¿Es fácil familiarizarse con la herramienta y aprender a usarla?
- *Eficiencia*: ¿Pueden los usuarios resolver sus tareas sin un esfuerzo innecesario? ¿Reacciona rápido?
- *Confianza*: ¿El usuario se siente en control de la interacción? ¿Es seguro y predecible?
- *Estímulo*: ¿Es emocionante y motivador utilizar la herramienta? Y
- *Novedad*: ¿Es creativo el diseño de la herramienta? ¿Capta el interés de los usuarios?

Además de la evaluación del sistema como un todo, se evaluaron también las visualizaciones de manera individual utilizando los siguientes criterios:

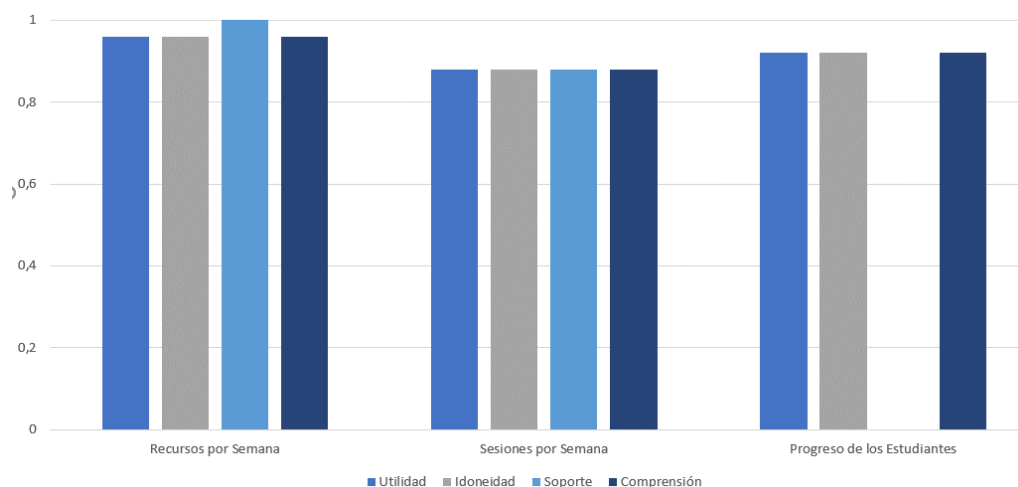
- *Utilidad*: ¿La información presentada en la visualización es relevante para el evaluador?
- *Idoneidad*: ¿Las visualizaciones son adecuadas para la información presentada?
- *Soporte*: ¿El tooltip de las visualizaciones poseen información extra útil?
- *Comprensión*: ¿El evaluador entendió el objetivo del gráfico?

6.6.1. Evaluación Vista Docente

El cuestionario para la evaluación de los docentes fue estructurado en ocho partes, la primera con una pregunta introductoria para conocer si el evaluador tenía experiencia previa usando dashboards de visualizaciones. Desde la segunda hasta la séptima parte se enfocó en evaluar cada visualización de la vista del Docente de la herramienta, haciendo cuatro preguntas respecto a cada visualización, cada una de ellas enfocadas a cubrir los criterios antes mencionado (Utilidad, Idoneidad, Soporte, Comprensión - Figura 112 hasta Figura 117). Por último, la octava parte contempló preguntas para la evaluación general de la herramienta, cada una enfocada a cubrir un criterio de UEQ (Figura 118).

Figura 112

Criterios de Usabilidad para la Evaluación de las Visualizaciones de la Sección Indicadores Generales (Vista Docente)

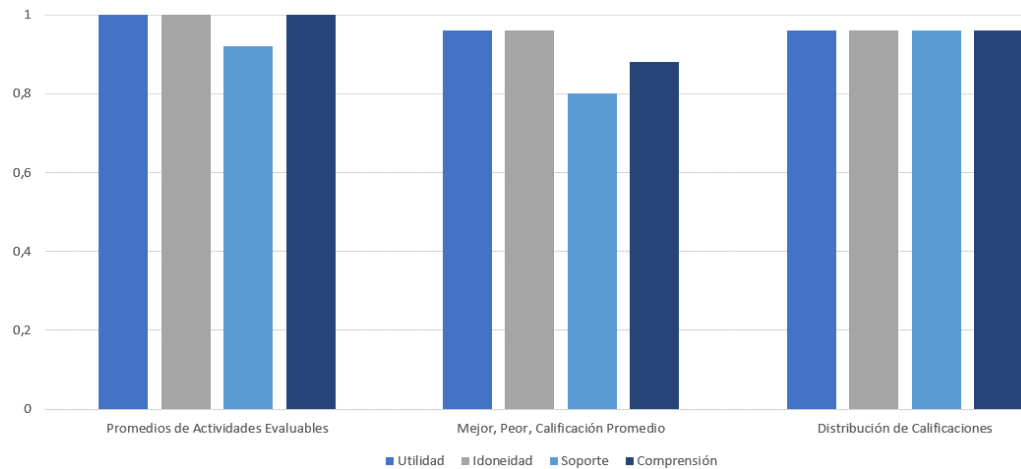


Nota. Elaboración Propia.



Figura 114

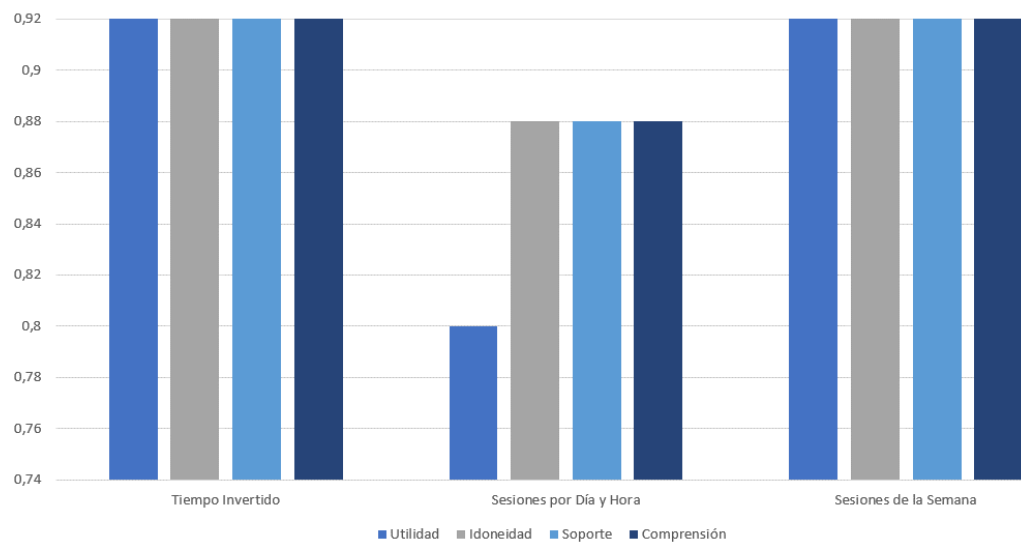
Criterios de Usabilidad para la Evaluación de las Visualizaciones de la Sección Seguimiento de Calificaciones (Vista Docente)



Nota. Elaboración Propia.

Figura 115

Criterios de Usabilidad para la Evaluación de las Visualizaciones de la Sección Sesiones de Estudio (Vista Docente)



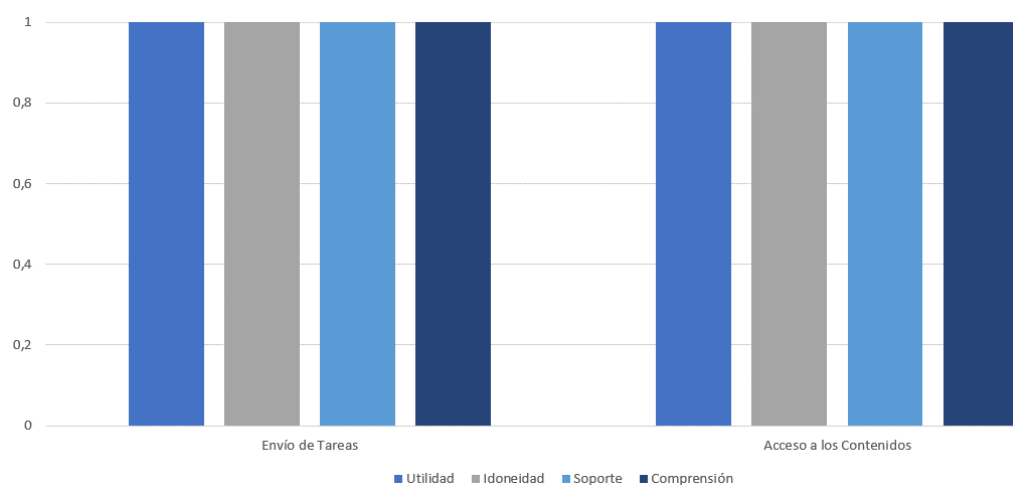
Nota. Elaboración Propia.



Figura 115

Criterios de Usabilidad para la Evaluación de las Visualizaciones de la Sección Seguimiento de Tareas

(Vista Docente)

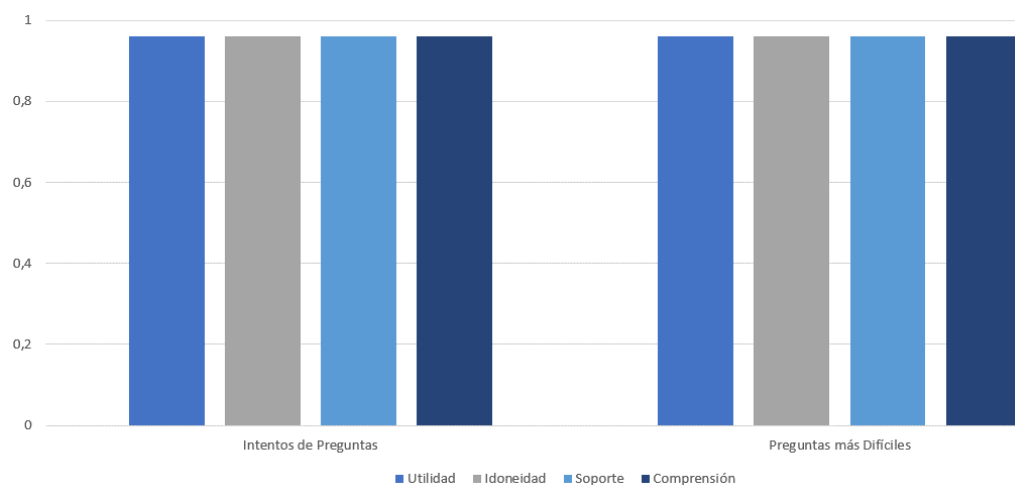


Nota. Elaboración Propia.



Figura 116

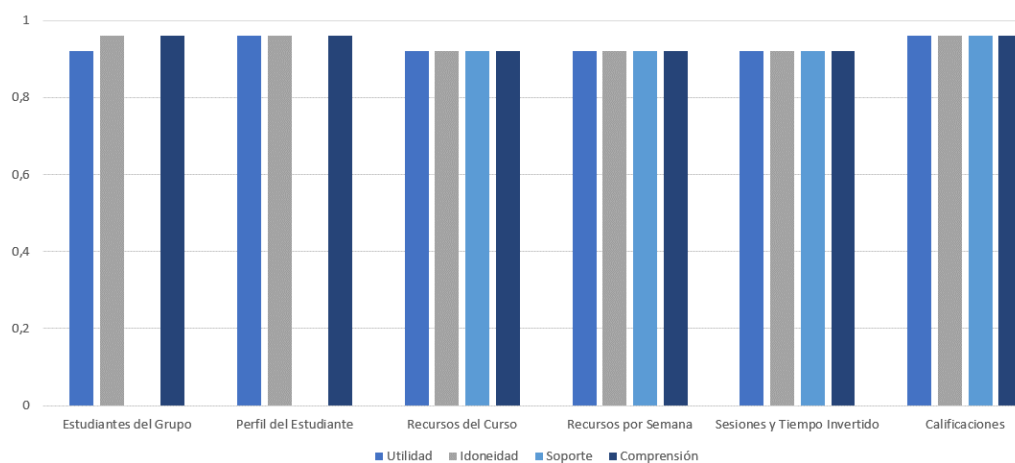
Criterios de Usabilidad para la Evaluación de las Visualizaciones de la Sección Seguimiento de Evaluaciones (Vista Docente)



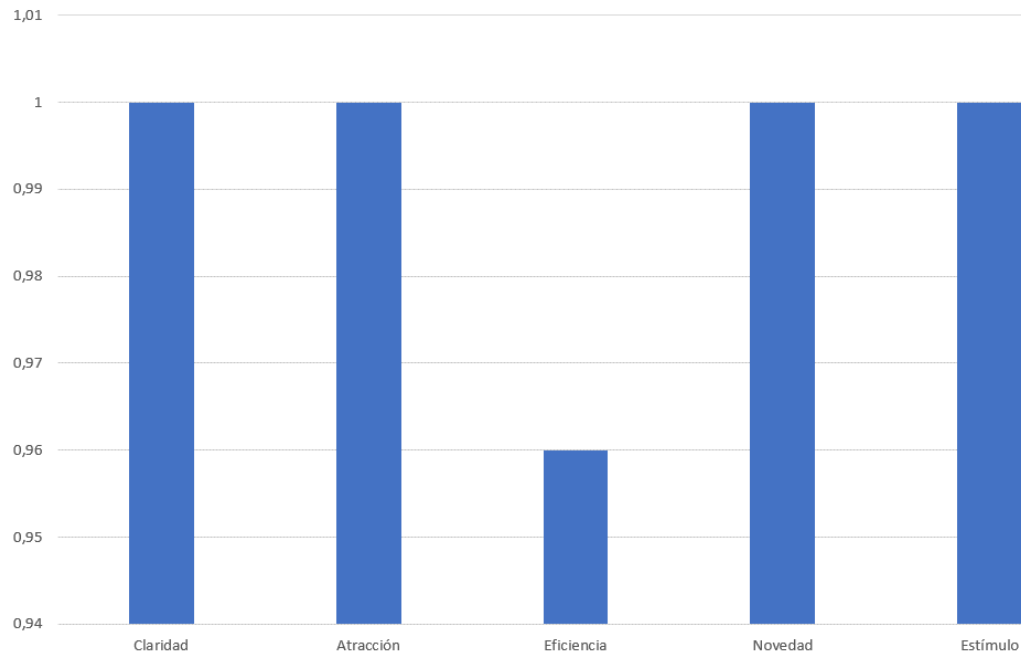
Nota. Elaboración Propia.

Figura 117

Criterios de Usabilidad para la Evaluación de las Visualizaciones de la Sección Deserción (Vista Docente)



Nota. Elaboración Propia.

Figura 118*Criterios de Usabilidad para la Evaluación de la Herramienta en General (Vista Docente)*

Nota. Elaboración Propia.

Los cinco evaluadores que participaron tenían experiencia previa usando dashboards, por lo que se esperaba que ninguno tenga problemas de usabilidad.

En la sección Indicadores Generales (Figura 112), se observa que los puntajes para las visualizaciones son considerablemente altos (cerca de 1), lo que llevaría a pensar que la información en esta sección completa es de mucha utilidad para los docentes. Una particularidad de esta figura se observa en los puntajes de la visualización *Progreso de los Estudiantes*, en el cual el valor de Soporte es cero. Esto se debe a que esta visualización no poseía un tooltip, por lo que en las evaluaciones no hubo pregunta sobre este criterio.

Las visualizaciones de la sección Sesiones de Estudio (Figura 113) obtuvieron considerablemente relativamente altos para las visualizaciones *Tiempo Invertido* y *Sesiones de la Semana*. Estos puntajes demuestran que estas visualizaciones son muy útiles para los evaluadores y no necesitan una mejora considerable. Por otro lado, la visualización *Sesiones por Día y Hora* presenta los puntajes más bajos para los criterios cuatro criterios analizados y especialmente en Utilidad. Estos puntajes implicarían una mejora a esta visualización, considerando una reestructuración

completa de la información de manera que sea más útil para los docentes y mostrada con un gráfico más adecuado.

Para las visualizaciones de las secciones Seguimiento de Tareas (Figura 115) y Seguimiento Evaluaciones (Figura 116) se observa los puntajes más altos de toda la evaluación. Al igual que la sección Indicadores Generales, todos los gráficos de esta sección son muy relevantes para los docentes y la información que aportan es muy útil.

En la sección Seguimiento de Calificaciones (Figura 114) se observa que la visualización susceptible a una es *Mejor, Peor y Calificación Promedio*. En esta visualización se podría considerar un rediseño del tooltip de ayuda y una explicación adicional de lo que muestra el gráfico mediante un mensaje informativo. De la misma manera, el rediseño del tooltip podría considerarse en la visualización *Promedios de Actividades Evaluables*, ya que el puntaje en el criterio Soporte está por debajo de la media.

Para la sección de Deserción (Figura 117) se observa un comportamiento similar a las anteriores, los puntajes son considerablemente altos, lo que implicaría una gran utilidad de las visualizaciones de esta sección para el docente. Una particularidad en los puntajes de esta sección es la ausencia de calificación para el criterio Soporte en las visualizaciones *Estudiantes del Grupo* y *Perfil del Estudiante*. Esto se debe a que no hubo una pregunta relacionada a Soporte para estas visualizaciones.

Con los criterios para evaluar la herramienta de forma general se observa que los puntajes obtenidos para Claridad, Atracción, Eficiencia, Novedad y Estímulo oscilan entre 0,96 y 1. Este puntaje refleja que el uso general de la herramienta resultó satisfactorio para los docentes. Se puede inferir que la herramienta fue fácil de usar, aprender, además permite a los usuarios interpretar las visualizaciones sin ningún esfuerzo adicional y causa interés en el usuario, lo que se esperaría ya que todos los evaluadores tenían experiencia previa utilizando dashboards.

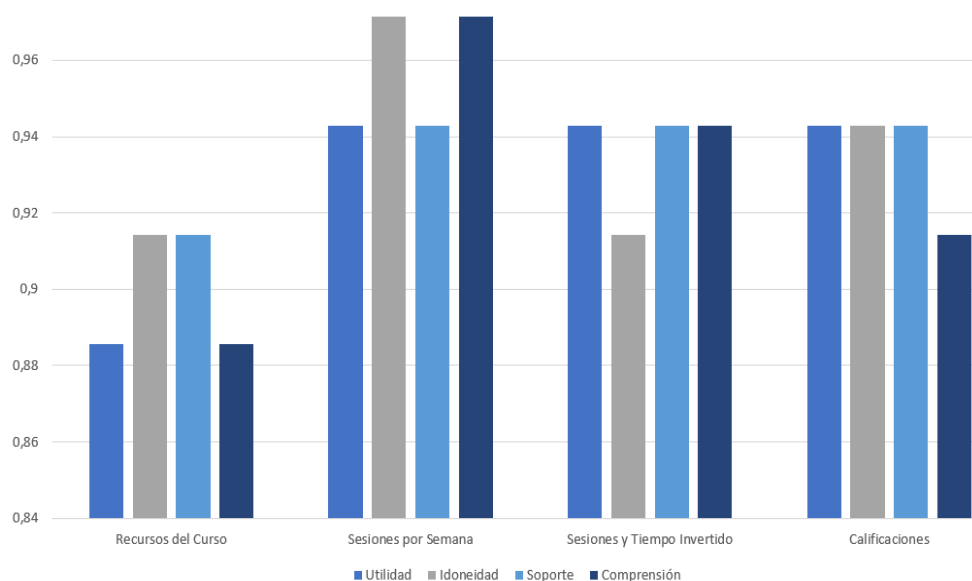
6.6.2. Evaluación Vista Estudiante

El cuestionario para la evaluación de los estudiantes fue estructurado en tres partes, la primera con una pregunta introductoria para conocer si el evaluador tenía experiencia previa usando dashboards de visualizaciones. La segunda parte se enfocó en evaluar cada visualización de la vista del Estudiante de la herramienta, haciendo cuatro preguntas respecto a cada visualización, cada una de ellas enfocadas a cubrir los criterios antes mencionado (Utilidad, Idoneidad, Soporte, Comprensión - Figura 119, Figura 120). Por último, la tercera parte contempló preguntas para la evaluación general de la herramienta, cada una enfocada a cubrir un criterio de UEQ (Figura 121).



Figura 119

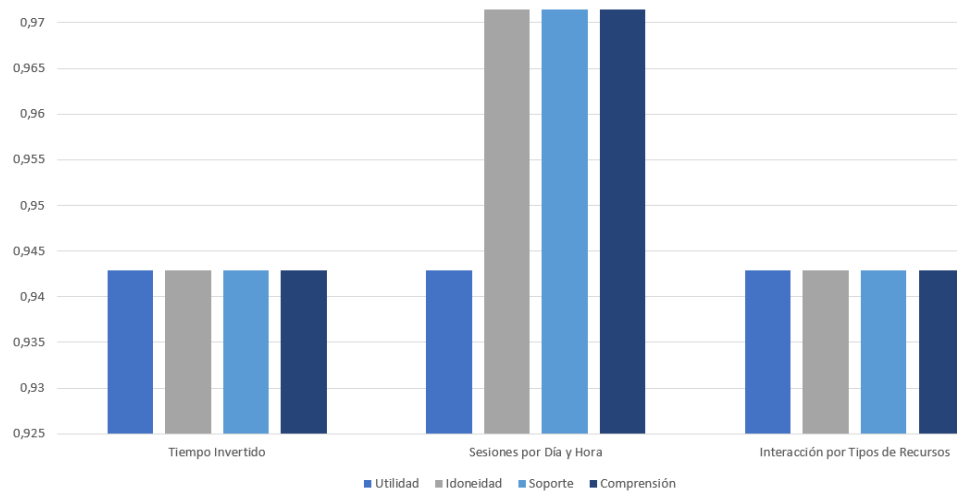
Criterios de Usabilidad para la Evaluación de las Visualizaciones de la Sección Indicadores Generales (Vista Estudiante)



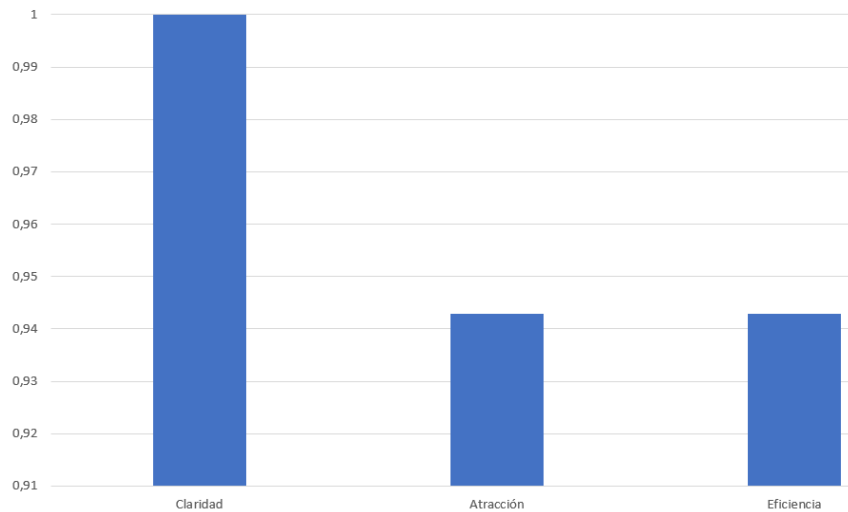
Nota. Elaboración Propia.

Figura 120

Criterios de Usabilidad para la Evaluación de las Visualizaciones de la Sección Sesiones de Estudio (Vista Estudiante)



Nota. Elaboración Propia.

Figura 121*Criterios de Usabilidad para la Evaluación de la Herramienta en General (Vista Estudiante)*

Nota. Elaboración Propia.

De los siete evaluadores que participaron, cuatro (53%) tenían experiencia previa usando dashboards mientras que tres evaluadores (42.9%) no habían usado dashboards antes. Si bien no se cuenta con una cantidad de evaluadores considerable, la cantidad de participantes que se tiene está equilibrada y se esperaría que la mitad de evaluadores sin experiencia revelen información útil sobre la usabilidad de la herramienta para personas sin experiencia previa.

En la sección Indicadores Generales (Figura 119), se observa que los puntajes para las visualizaciones oscilaron entre 0.88 y 0.98. Para la visualización *Recursos del Curso*, en relación a las demás visualizaciones de la sección, se tiene los puntajes más bajos para los criterios Utilidad, Idoneidad, Soporte y Comprensión. Estos puntajes revelan que en general esta visualización es la que menos información aporta a los evaluadores y podría considerarse una mejora. La visualización *Sesiones por Semana* es la que presenta los puntajes más altos para los criterios Idoneidad y Comprensión, mientras que los puntajes para Utilidad y Soporte están dentro de la media. Finalmente, las visualizaciones *Sesiones y Tiempo Invertido* y *Calificaciones* presentan gráficos similares y dentro de la media para los criterios Utilidad y Soporte. Para la visualización *Sesiones y Tiempo Invertido*, su bajo puntaje en el criterio Idoneidad podría sugerir el utilizar un diferente gráfico para presentar la información, mientras que para la visualización *Calificaciones* su bajo puntaje para el criterio Comprensión podría motivar a presentar la información de forma distinta.



Las visualizaciones de la sección Sesiones de Estudio (Figura 120) obtuvieron puntajes que oscilan entre 0.94 y 0.98. Estos puntajes demuestran que las visualizaciones de esta sección del plugin son muy útiles para los evaluadores y no necesitan una mejora considerable. En el caso de las visualizaciones *Tiempo Invertido* e *Interacción por Tipos de Recursos*, éstas presentan puntajes superiores a 0.9 para todos los criterios de evaluación, los cuales a pesar de ser los más bajos, superan el 90% de la valoración total, dando a entender su importancia y utilidad. En el caso de la visualización *Sesiones por Día y Hora*, ésta presenta los puntajes más altos para los criterios Idoneidad, Soporte y Comprensión, mientras que para el criterio Utilidad, el puntaje está dentro de la media de las demás visualizaciones.

Con los criterios para evaluar la herramienta de forma general (Figura 121), se observa que los puntajes obtenidos para Claridad, Atracción y Eficiencia oscilan entre 0,91 y 1. Este puntaje refleja que el uso general de la herramienta resultó satisfactorio para los evaluadores. Se puede inferir que la herramienta fue fácil de usar, aprender, además permite a los usuarios interpretar las visualizaciones sin ningún esfuerzo adicional y causa interés en el usuario. Además, se ha visto que la premisa de la mitad de evaluadores sin experiencia no afectó en nada las los puntajes altos de los criterios de las visualizaciones, lo que también da una idea de lo usable que es la herramienta. Sin embargo, también es necesario aclarar que la cantidad reducida de evaluadores podría haber afectado los puntajes, puesto que no se tiene un amplio rango de puntajes que puedan confirmar lo inferido. Se necesitaría de una cantidad mayor de personas que participen del proceso de evaluación para lograr datos más exactos sobre la usabilidad de la herramienta. A pesar de ello, la evaluación presentada con los datos obtenidos resultó muy útil y da una buena idea de la usabilidad de la herramienta.



Capítulo 7: Conclusiones

7.1. Conclusiones

Este trabajo de tesis presenta FlipMyLearning, una herramienta de monitoreo del proceso de aprendizaje del estudiante, así como también de su comportamiento, por medio del uso de dashboards de visualizaciones de datos y dashboards de predicción de abandono de los estudiantes en la plataforma Moodle. FML permite tanto a docentes como a estudiantes monitorear el proceso de aprendizaje en un curso sobre la plataforma Moodle. De los resultados obtenidos y en base a las preguntas de investigación planteadas, en el capítulo uno, se puede concluir que:

(1) Con relación a la primera pregunta de investigación ¿Qué secuencias de aprendizaje se pueden descubrir a partir de las interacciones de los estudiantes en un curso virtual de Moodle?; en esta tesis se desarrolló un análisis exploratorio del comportamiento de los estudiantes utilizando técnicas de machine learning para descubrir secuencias de aprendizaje basado en las interacciones de los estudiantes con la plataforma. El análisis del comportamiento de los estudiantes mediante Minería de Procesos permitió generar mapas de procesos que sirvieron para identificar el comportamiento de los estudiantes en cada semana de un curso en relación con la sección asignada a cada semana. Se esperaba que los estudiantes interactúen cada semana con los contenidos del curso asignados a tal semana, pero como se observó en los análisis, los estudiantes empiezan el curso sin tener un comportamiento esperado, luego normalizan sus actividades en el transcurso de las semanas hasta antes del final del curso, en donde el comportamiento vuelve a diferir del esperado. Este comportamiento podría tener muchas causas. Al inicio, los estudiantes no tienen un comportamiento esperado ya que les cuesta tomar el ritmo del curso de manera inicial, luego en el transcurso adoptan el ritmo de trabajo adecuado, lo que les permite avanzar en el curso sin mayor contratiempo, y al final del curso, el desgaste y la acumulación de actividades hacen que el comportamiento vuelva a diferir de lo que se esperaba. La elección de utilizar Minería de Procesos y no algoritmos de Machine Learning como tal para la detección del comportamiento tuvo su razón en que, si bien varios estudios sugieren que se puede analizar el comportamiento de los estudiantes con técnicas de Machine Learning, el uso de la Minería de Procesos presenta una utilidad potencial para el análisis de registros de eventos que registran tanto estudiantes como docentes en un entorno de aprendizaje virtual (Juhaňák, Zounek, & Rohlíková, 2019).



(2) Con relación a la segunda pregunta de investigación ¿Qué variables son útiles para predecir el abandono de estudiantes en un curso virtual de Moodle?; en esta tesis se identificaron las variables más importantes para predecir el abandono de estudiantes. Las variables utilizadas como predictores para construir el modelo predictivo se calcularon en base a tres enfoques: el primero fue el tiempo invertido por el estudiante en el curso. El segundo enfoque consideró la interacción de los estudiantes con los diversos recursos y actividades de diseño de aprendizaje disponibles en el curso. Y el tercer enfoque incluyó las interacciones del estudiante con sus compañeros del curso y con el docente. Estos tres enfoques permitieron generar indicadores en base al tiempo invertido (sesiones de estudio), la profundidad cognitiva (interacciones con recursos) y la amplitud social (interacciones entre estudiantes) de las actividades y recursos de Moodle. El enfoque del Tiempo Invertido permitió generar tres indicadores: el Número de Sesiones de estudio, el Tiempo Invertido en el curso y los Días Activos del estudiante en la plataforma. Para determinar el número de sesiones se ha hecho uso de la definición de sesión. Para esta definición se han considerado períodos de inactividad entre las interacciones de los estudiantes dentro de un curso de Moodle. Eso significa que cualquier par de interacciones cuya diferencia en el tiempo en el que ocurrió es menor a 30 minutos se consideran parte de la misma sesión. De otra forma, si el tiempo entre dos interacciones es mayor a 30 minutos, se consideran diferentes sesiones. El tiempo invertido representa la cantidad de tiempo en minutos que el estudiante permaneció activo en la plataforma. Esta variable se calcula sumando la duración de cada sesión de estudio que realizó en el curso. Los días activos del estudiante representan la cantidad de días que el estudiante realizó al menos una sesión de estudio desde que inició el curso.

Los indicadores relacionados a la profundidad cognitiva se definieron en base al tipo de actividad o recurso de Moodle y en qué medida el alumno demuestra un compromiso cognitivo en esa actividad. Este compromiso se mide en niveles que varían de 0 a 5, donde 0 indica que el alumno ni siquiera ha visto la actividad. Los niveles de profundidad cognitiva potencial son: el estudiante ha visto los detalles de la actividad (0 o 100%), el estudiante ha enviado contenido a la actividad (0, 50 o 100%), el estudiante ha visto los comentarios de un instructor o un compañero para la actividad (0, 33, 67, 100%), el estudiante ha proporcionado comentarios al instructor o un compañero dentro de la actividad (0, 25, 50, 75 o 100%) y el estudiante ha revisado y/o vuelto a enviar contenido a la actividad (0, 20, 40, 60, 80 o 100%).

Los indicadores relacionados a la amplitud social miden la capacidad de los participantes para identificarse o comunicarse con el grupo o curso de estudio. El nivel de amplitud varía de 0 a 2, donde 0 indica que el alumno no ha interactuado con nadie. Los niveles de amplitud social: el



estudiante no ha interactuado con ningún otro participante en esta actividad (0 o 100%) y el estudiante ha interactuado con al menos otro participante (0, 50 o 100%).

De esta manera, la profundidad cognitiva y la amplitud social se define a nivel de cada contenido individual del curso, por lo que la cantidad de indicadores será igual al doble del número de contenidos que tenga un curso. Es necesario mencionar que entre los diversos tipos de contenidos que permite crear Moodle existe el recurso Etiqueta (label), el cual no fue considerado para obtener indicadores de Profundidad Cognitiva y Amplitud Social. Ya que este tipo de recurso siempre presentaba el mismo valor para cada uno de los estudiantes de un curso. Además, no se tomaron en cuenta los recursos tipo módulos externos a los de una instalación estándar de Moodle. Teniendo en cuenta la generación de indicadores y variables predictoras, el modelo de datos para la predicción se conforma de una matriz conformada por filas y columnas. La cantidad de filas corresponde a la cantidad de alumnos matriculados en el curso, mientras que, la cantidad de columnas depende de la cantidad de módulos, (recursos y actividades del curso), número de sesiones, tiempo invertido, cantidad de días que el estudiante estuvo activo en la plataforma y el identificador del estudiante. Los indicadores se calcularon en base a las fechas de inicio y finalización del curso. En los casos de cursos que no tienen una fecha de finalización configurada, los indicadores se generaron con todos los datos disponibles desde la fecha de inicio del curso.

(3) Con relación a la tercera pregunta de investigación ¿Cómo implementar un dashboard para visualizar el comportamiento de estudiantes en un curso virtual de Moodle?; en esta tesis se implementó FlipMyLearning, un dashboard en Moodle para visualizar el comportamiento de los estudiantes en un curso virtual. Lo más importante en el diseño y desarrollo de FlipMyLearning fue proveer a los docentes de una herramienta con la capacidad de saber qué realmente está sucediendo en el transcurso de un curso de manera que sea posible implementar cambios tempranos para alinear los comportamientos estudiantiles a los objetivos académicos. FlipMyLearning ha sido construido para la plataforma Moodle 3.9+, es instalable y escalable, ya que está desarrollado bajo las directrices arquitectónicas modulares de Moodle y no contiene dependencias de otros módulos. FlipMyLearning muestra información, de manera gráfica e intuitiva, sobre el proceso de aprendizaje de los estudiantes con un gran número de gráficas divididas en secciones fácilmente identificables como son: configurar semanas, sesiones de estudio, seguimiento de tiempo, seguimiento de tareas, seguimiento de calificaciones, seguimiento de evaluaciones y deserción, secciones que han sido explicadas en la sección de [Funcionalidades](#) del Capítulo 6. El proceso de desarrollo de FlipMyLearning estuvo dirigido por una metodología híbrida compuesta por cinco etapas: Recolección, Procesamiento,



Análisis, Visualización y Evaluación. Las actividades más importantes para la implementación de este dashboard fueron la recopilación, preparación y transformación de los datos, así como el análisis de estos en base a interacciones entre los diferentes actores del proceso educativo (docente, estudiante) con la plataforma Moodle. Estas interacciones consisten en actividades evaluables del curso, acceso a recursos, tiempo empleado y sesiones de estudio a lo largo de diferentes semanas de estudio. Uno de los aspectos que significó mayor complejidad para el desarrollo del dashboard fue la comprensión del modelo de datos sobre el que se basa Moodle debido a la cantidad de tablas de datos (461 en la versión de Moodle 3.9) y relaciones entre las tablas de datos. Por otro lado, la falta de documentación sobre la implementación de extensiones o plugins para esta plataforma también significó un desafío ya que no existen guías con la suficiente información para introducir al desarrollo de plugins para Moodle a personas inexpertas.

(4) En esta tesis se evaluó la usabilidad del dashboard con docentes y estudiantes en un contexto real. El proceso como tal se llevó a cabo mediante cuestionario elaborados en Google Forms y aplicados a cinco docentes y siete estudiantes. A pesar de que no se contó con un gran número de evaluadores, los resultados obtenidos dan una buena idea de la utilidad de la herramienta tanto para docentes como para estudiantes. Las evaluaciones se realizaron tomando como referencia criterios de usabilidad para cada visualización del dashboard y para la herramienta como un todo. En general, se obtuvieron puntajes que oscilaron entre 0.8 y 1 de Utilidad de la información, Idoneidad del Gráfico, ayuda de Soporte y Comprensión de la Información para cada una de las visualizaciones. Así mismo, los puntajes de la evaluación a la herramienta en general oscilaron entre 0.94 y 1 en la escala EQU, lo que dice que el sistema desarrollado es agradable a los usuarios (Atractivo), es fácil de aprender a usar (Claridad), se puede usar sin conocimientos previos (Eficiencia), motiva a volver a usarse (Estímulo) y llama la atención del usuario (Novedad). Además, los evaluadores sugirieron que se agreguen indicadores como la cantidad de estudiantes en riesgo de deserción y la cantidad de estudiantes en riesgo de abandono para el docente y recursos a los que se recurren mayormente, cantidad de actividades calificables y estrategias para auto y coevaluación para el estudiante.

Finalmente, en los últimos años ha habido un considerable interés sobre LA en cursos virtuales de aprendizaje, y el potencial de recopilar "grandes datos" de los estudiantes que usan estas plataformas. Sin embargo, las plataformas actuales tienen una capacidad limitada para proporcionar datos y visualizaciones que pueden usarse fácilmente para ayudar a los docentes, diseñadores de cursos y estudiantes en la toma de decisiones. En esta tesis se ha demostrado que, con un análisis



exhaustivo de los datos proporcionados por Moodle, es posible generar visualizaciones y diagnósticos asociados con los datos analíticos de aprendizaje. De hecho, el verdadero valor de la analítica de aprendizaje para los LMS, y para el aprendizaje en línea, o combinado en general, solo se verá cuando las plataformas de aprendizaje virtual incorporen herramientas de análisis y visualizaciones similares a las presentadas en este documento para que los usuarios finales accedan de manera rutinaria y fácil.

7.2. Recomendaciones

Las siguientes recomendaciones realizadas están estrictamente dirigidas al docente, esto con el afán de que FlipMyLearning sea usada de la mejor manera aprovechando al máximo todas sus capacidades.

- Se recomienda al docente crear una estructura adecuada del curso. Moodle permite definir esta estructura en base a una planificación previa. Esta planificación se denomina *Configuración del Libro de Calificaciones* donde se definen secciones y puntuaciones/ponderaciones a las diferentes actividades. Esta configuración permitirá obtener mejores resultados en gráficos de la sección [Seguimiento de Calificaciones](#). Hay que asegurar que el plan del curso esté bien diseñado y que la alineación entre los resultados de aprendizaje planificados y las diferentes actividades propuestas en el curso mantengan correlación permite garantizar que el resultado esté alineado con lo que los estudiantes experimentan y aprenden. Puesto que, el diseño del curso es uno de los procesos importantes que afectan la efectividad de cualquier herramienta analítica.
- Se recomienda al docente configurar las opciones de finalización de módulos de manera personalizada para cada recurso del curso. La finalización de actividad (*Activity Completion en inglés*) le permite al docente configurar criterios de finalización dentro de la configuración de una actividad específica. Dependiendo del tipo de actividad, hay diferentes requisitos para la finalización. Por ejemplo, una página puede requerir ser *vista*; un examen puede tener el requisito de *tener una calificación asignada*, un foro puede requerir *publicaciones discusiones o respuestas*. Cuando un estudiante finalice una actividad en base a las condiciones configuradas, una casilla de verificación aparece junto a la actividad. Esta configuración permite a FlipMyLearning mejorar la apreciación de progreso de los estudiantes en el curso, como se muestra en las secciones de [Sesiones de Estudio](#) y [Deserción](#).



- Se recomienda al docente analizar los datos generados en la Sección Deserción del plugin luego de finalizar una semana de trabajo (Lunes a Viernes). El plugin está configurado para recalcular los grupos de estudiantes cada viernes mientras dure el curso.

7.3. Trabajos Futuros

Los análisis y visualizaciones de los datos analíticos de aprendizaje presentados en este estudio contribuyen a proporcionar una mayor comprensión de la actividad de los estudiantes en un curso Moodle basado en sus interacciones. El alcance del estudio está limitado al hallazgo, visualización de patrones de comportamiento de aprendizaje y predicción de abandono de los estudiantes en un curso virtual de la plataforma Moodle. Sin embargo, todavía hay investigación que se puede hacer en esta área. Por ejemplo:

(1) Centrarse en la aplicación de los conocimientos generados por este estudio en un proceso de rediseño del material educativo. Los archivos de datos generados por las gráficas gracias a HighCharts en los formatos de CSV, XLS como se mencionó en la sección de [Funciones Adicionales](#), pueden servir como insumos para nuevos análisis.

(2) Explorar otras técnicas de minería de datos y Machine Learning para mejorar la predicción de estudiantes en el curso. Generar un conjunto de datos etiquetados con resultados de predicciones de FlipMyLearning en diferentes cursos para crear un conjunto de datos de entrenamiento para futuras predicciones basadas en aprendizaje supervisado.

(3) Incorporar nuevas visualizaciones que reflejan las interacciones de los estudiantes en el curso. Estudiar más a detalle las sesiones que realizan los estudiantes, con nuevos parámetros que logren determinar cuáles son las sesiones que mejor se aprovechan, que tiempo se emplea y por qué. La investigación adicional también podría examinar el concepto de motivación de los estudiantes y la relación que esto tiene con los patrones de participación y rendimiento en cursos en líneas.

(4) Desarrollar nuevas características para Moodle que permitan realizar análisis de datos con información más precisa. Una funcionalidad muy necesaria es un nuevo sistema de registro de eventos para Moodle, el cual permita hacer un seguimiento más exacto de las interacciones que pertenecen a una misma sesión de un usuario.

(5) Desarrollar un mecanismo que monitoree de mejor manera las interacciones del estudiante, de tal forma que sea posible almacenar en el registro de eventos de Moodle el momento



en que un estudiante cierra el sitio web, lo que podría considerarse como la finalización de una sesión de estudio.

En la actualidad existen muchos estudios sobre análisis de datos en plataformas LMS como Coursera u otras. En dichos estudios, se recalca la importancia de analizar interacciones con video lecturas tales como: iniciar una video lectura, reiniciar una video lectura, finalizar una video lectura, repetir una video lectura. En la versión 3.9 (más actual a la fecha), Moodle posee un componente para reproducir videos en un curso, sin embargo, las interacciones antes descritas no son registradas en el log de eventos, lo cual deja la imposibilidad de analizar las interacciones con video lecturas y también representa una de las principales razones por las que no existen muchos estudios de análisis de datos sobre cursos online de Moodle.

Finalmente, el proceso de evaluación de FlipMyLearning dejó claro la utilidad de la herramienta para los docentes y estudiantes. Futuras investigaciones podrían enfocarse en desarrollar versiones del software para otros LMS.



Bibliografía

- Abad León, F. L., & Vizhñay Enderica, E. D. (12 de Octubre de 2020). Microservicios aplicados a la integración y análisis de datos orientados al cálculo de una tarifa diferencial para aparcaderos de vehículos privados. Caso de estudio: Campus Central de la Universidad de Cuenca. *Tesis de Pregrado*, 25-41. Cuenca, Azuay, Ecuador.
- Abeysekera, L., & Dawson, P. (2014). Motivation and cognitive load in the flipped classroom: definition, rationale and a call for research. *Higher Education Research & Development*, 1-14. doi:<http://dx.doi.org/10.1080/07294360.2014.934336>
- Abril Lara, C. (Febrero de 2016). Herramienta Business intelligence aplicando la metodología Hefesto v2.0 para generar reportes estadísticos de las emergencias atendidas en el SIS ECU911 Zona 3. *Tesis de pregrado*, 6-19. Ambato, Tungurahua, Ecuador. Obtenido de http://repositorio.uta.edu.ec/bitstream/123456789/23659/1/Tesis_t1148si.pdf
- Abubakar, R., Kamsin, A., & Abdullah, A. (2020). Challenges in the online component of blended learning: A systematic review. *Computers & Education*. doi:<https://doi.org/10.1016/j.compedu.2019.103701>
- Agarwal, S. (2013). Data Mining: Data Mining Concepts and Techniques. *2013 International Conference on Machine Intelligence and Research Advancement*, (págs. 203-207). Katra. doi:10.1109/ICMIRA.2013.45
- Aguilar, J. (15 de Octubre de 2019). ¿Qué es el patrón MVC en programación y por qué es útil? Obtenido de campusmvp: <https://www.campusmvp.es/recursos/post/que-es-el-patron-mvc-en-programacion-y-por-que-es-util.aspx>
- Al-Ajlan, A., & Zedan, H. (2008). Why Moodle. *2008 12th IEEE International Workshop on Future Trends of Distributed Computing Systems*, (págs. 58-64). Kunming. doi:10.1109/FTDCS.2008.22
- Aldowaha, H., Al-Samarraiea, H., & Mohamad Fauzyb, W. (2019). Educational data mining and learning analytics for 21st century higher education: A review and synthesis. *Telematics and Informatics*, XXXVII, 13-49. doi:<https://doi.org/10.1016/j.tele.2019.01.007>
- Anthony, J. (2018). *Top 10 Alternatives to Moodle: Popular LMS Solutions To Consider*. Obtenido de financesonline: <https://financesonline.com/top-10-alternatives-to-moodle-popular-lms-solutions-to-consider/>
- Arkadiusz, K. (2020). *PHP-ML - Machine Learning library for PHP*. Obtenido de php-ml.readthedocs.io: <https://php-ml.readthedocs.io/en/latest/>



- Bakharia, A., & Dawson, S. (2011). SNAPP: a bird's-eye view of temporal participant interaction. *Proceedings of the 1st International Conference on Learning Analytics and Knowledge*, (págs. 168-173). doi:<https://doi.org/10.1145/2090116.2090144>
- Barrios, W., Fernández, M., Godoy, M., & Mariño, S. (2012). De Moodle a Entornos Personales de Aprendizaje (PLE): Introducción de herramientas sociales a una plataforma e-learning. *Simposio sobre la Sociedad de la Información*, (págs. 93-104). Obtenido de http://41jaiio.sadio.org.ar/sites/default/files/8_SSI_2012.pdf
- Bernabeu, D. (6 de Mayo de 2009). *dataprix*. Obtenido de <https://www.dataprix.com/es/data-warehousing-y-metodologia-hefesto/data-warehousing-y-metodologia-hefesto>
- Bogarín, A., Cerezo, R., & Romero, C. (Septiembre de 2017). A survey on educational process mining. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, 8. doi:<https://doi.org/10.1002/widm.1230>
- Bogarín, A., Romero, C., Cerezo, R., & Sánchez Santillán, M. (2014). Clustering for improving Educational Process Mining. *Proceedings of the Fourth International Conference on Learning Analytics And Knowledge*, (págs. 11-15). doi:<https://doi.org/10.1145/2567574.2567604>
- Büchner, A. (2016). *Moodle 3 Administration. An administrator's guide to configuring, securing, customizing, and extending Moodle*. Birmingham: Packt Publishing Ltd. Obtenido de <https://www.packtpub.com/product/moodle-3-administration-third-edition/9781783289714>
- Buenaño Fernández, D., & Luján Mora, S. (2016). Exploring approaches to educational data mining and learning analytics, to measure the level of acquisition of student's learning outcome. *Conference: International Conference on Education and New Learning Technologies*, (págs. 1845-1850). doi:10.21125/edulearn.2016.1368
- Bustos Barrera, S., & Mosquera Artieda, V. (Marzo de 2013). Análisis, diseño e implementación de una solución de BI para la generación de indicadores y control de desempeño, en la empresa OTECEL, utilizando la metodología Hefesto v2.0. *Tesis de pregrado*. Ecuador. Obtenido de <http://repositorio.espe.edu.ec/xmlui/bitstream/handle/21000/6305/T-ESPE-047033.pdf?sequence=1&isAllowed=y>
- Calvanese, D., Giacomo, G., Lenzereni, M., Nardi, D., & Rosati, R. (2001). Data integration in Data Warehousing. *International Journal of Cooperative Information Systems*, 10(3), 237-271. doi:<https://doi.org/10.1142/S0218843001000345>
- Castañeda Quintero, L., & López Vicent, P. (2007). Entornos virtuales de enseñanza aprendizaje libres: Moodle para profesores. *Herramientas Telemáticas Para La Enseñanza Universitaria*



- En El Marco Del Espacio Europeo De Educación Superior*. Obtenido de <https://digitum.um.es/digitum/bitstream/10201/13417/1/moodle.pdf>
- Chakraborty, S., & Nagwani, N. (2014). Analysis and Study of Incremental DBSCAN Clustering Algorithm. *International Journal of Enterprise Computing and Business*. Obtenido de <https://arxiv.org/abs/1406.4754>
- Chen, M., Ebert, D., Hagen, H., Laramée, R., Liere, R., Ma, K., . . . Silver, D. (2009). Data, Information, and Knowledge. *IEEE Computer Graphics and Applications*, 29, págs. 12-19. doi:10.1109/MCG.2009.6
- Chung, M., & Gray, P. (1999). Special Section: Data Mining. *Journal of Management Information Systems*, 16(1), 11-16. doi:<https://doi.org/10.1080/07421222.1999.11518231>
- Coccia, M., & Watts, J. (2019). A theory of the evolution of technology: Technological parasitism and the implications for innovation management. *Journal of Engineering and Technology Management*, LV. doi:<https://doi.org/10.1016/j.jengtecman.2019.11.003>
- Cockrum, T. (2014). *Flipping your English class to reach all learners: Strategies and lesson plans*. New York, Estados Unidos. Obtenido de https://books.google.com.ec/books?hl=es&lr=&id=OIZiAgAAQBAJ&oi=fnd&pg=PP1&dq=Flipping+your+English+class+to+reach+all+learners:+Strategies+and+lesson+plans&ots=aUbA9aZwsi&sig=fqdDwfZ915cZ0vcnYKlqB1PNhGk&redir_esc=y
- Codish, D., Rabin, E., & Ravid, G. (2019). User behavior pattern detection in unstructured processes – a learning management system case study. *Interactive Learning Environments*, 699-725. doi:<https://doi.org/10.1080/10494820.2019.1610456>
- Coffrin, C., Corrin, L., de Barba, P., & Kennedy, G. (2014). Visualizing patterns of student engagement and performance in MOOCs. *Proceedings of the Fourth International Conference on Learning Analytics And Knowledge*, (págs. 83-92). Indiana. doi:<https://doi.org/10.1145/2567574.2567586>
- Costa, C., Alvelos, H., & Teixeira, L. (2012). The Use of Moodle e-learning Platform: A Study in a Portuguese University. *Procedia Technology*, V, 334-343. doi:<https://doi.org/10.1016/j.protcy.2012.09.037>
- D. Roberts, L., A. Howell, J., & Seaman, K. (2017). Give Me a Customizable Dashboard: Personalized Learning Analytics Dashboards in Higher Education. *Technology, Knowledge and Learning*, 317-333. doi:<https://doi.org/10.1007/s10758-017-9316-1>
- Den Beemt, A., Buijs, J., & Der Aalst, W. (2018). Analysing Structured Learning Behaviour in Massive Open Online Courses (MOOCs): An Approach Based on Process Mining and



- Clustering. *International Review of Research in Open and Distance Learning*, 39-60. doi:<https://doi.org/10.19173/irrodl.v19i5.3748>
- Drăgulescu, B., Bucos, M., & Vasiu, R. (2015). CVLA: Integrating Multiple Analytics Techniques in a Custom Moodle Report. *International Conference on Information and Software Technologies*, (págs. 115-126). doi:https://doi.org/10.1007/978-3-319-24770-0_11
- Dreizzen, E., Zangara, M. A., & Alonso, N. (2016). Proceso de implementación del nuevo entorno virtual de enseñanza y aprendizaje Moodle de la Facultad de Ciencias Médicas de la UNLP. *Planificación y temas críticos*, 17, 52-57. Obtenido de <http://sedici.unlp.edu.ar/handle/10915/54201>
- DualCube Team. (5 de Junio de 2020). *moodle.org*. Obtenido de Moodle Plugins: https://moodle.org/plugins/gradereport_quizanalytics
- Einhardt, L., Aires Tavares, T., & Cechinel, C. (2016). Moodle analytics dashboard: A learning analytics tool to visualize users interactions in moodle. *2016 XI Latin American Conference on Learning Objects and Technology (LACLO)*, (págs. 1-6). San Carlos. doi:10.1109/LACLO.2016.7751805
- Et-taleby, A., Boussetta, M., & Benslimane, M. (2020). Faults Detection for Photovoltaic Field Based on K-Means, Elbow, and Average Silhouette Techniques through the Segmentation of a Thermal Image. *International Journal of Photoenergy*. doi:<https://doi.org/10.1155/2020/6617597>
- Foley, É., & Guillemette, M. (2010). What is Business Intelligence? *International Journal of Business Intelligence Research*, 1-28. doi:10.4018/jbir.2010100101
- G. de Barba, P., Malekian, D., A. Oliveira, E., Bailey, J., Ryan, T., & Kennedy, G. (Marzo de 2020). The importance and meaning of session behaviour in a MOOC. *Computers and Education*, 146(103772). doi:10.1016/j.compedu.2019.103772
- García Alba, M. J. (2010). Análisis del desarrollo de extensiones para Moodle: Desarrollo de un módulo para la gestión de laboratorios docentes. *Tesis de Pregrado*. Alcalá de Henares, España.
- García, E., Romero, C., Ventura, S., & Castro, C. (2008). An architecture for making recommendations to courseware authors using association rule mining and collaborative filtering. *User Model User-Adap Inter*, XIX, págs. 99-132. doi:<https://doi.org/10.1007/s11257-008-9047-z>
- Garrison. (2009). Communities of Inquiry in Online Learning. *Encyclopedia of Distance Learning*, 352-355. doi:10.4018/978-1-60566-198-8.ch052



- Garrison, R., Anderson, T., & Archer, W. (2000). Critical Inquiry in a Text-Based Environment: Computer Conferencing in Higher Education. *The Internet and Higher Education*, II(3), 87-105. doi:[https://doi.org/10.1016/S1096-7516\(00\)00016-6](https://doi.org/10.1016/S1096-7516(00)00016-6)
- Gaudioso, E., & Talavera Méndez, L. J. (2005). Data mining to support tutoring in virtual learning communities: experiences and challenges. *Computer Science*. doi:10.2495/1-84564-152-3/12
- Gazzotti, A., Algieri, R., Doglioti, C., Mazzoglio, M., Rey, L., Jiménez Villarruel, H., . . . Pró, E. (5 de Mayo de 2011). Adhesión al espacio virtual de enseñanza y aprendizaje Moodle en alumnos de Anatomía. *Revista Argentina de Anatomía Online 2011*, II(3), 37-70. Obtenido de <https://core.ac.uk/download/pdf/25798041.pdf>
- Ginda, M., Suri, N., Bueckle, A., & Börner, K. (2016). Empowering Instructors in Learning Management Systems : Interactive Heat Map Analytics Dashboard. *Computer Science*. Obtenido de <https://www.semanticscholar.org/paper/Empowering-Instructors-in-Learning-Management-%3A-Map-Ginda-Suri/bc132c9bc131377a1e1d3809d181898d98791864>
- Golfarelli, M., Rizzi, S., & Cella, I. (2004). Beyond data warehousing: what's next in business intelligence? *Proceedings of the 7th ACM international workshop on Data warehousing and OLAP*, (págs. 1-6). doi:<https://doi.org/10.1145/1031763.1031765>
- Gómez Aguilar, D. A., Therón, R., & García Peñalvo, F. (2008). Understanding Educational Relationships in Moodle with ViMoodle. *2008 Eighth IEEE International Conference on Advanced Learning Technologies*, (págs. 954-956). Cantabria. doi:10.1109/ICALT.2008.276
- Gonzales Farro, G. D. (2018). Implementación de una solución de inteligencia de negocios utilizando la metodología Hefesto para las oficinas de contabilidad en universidades públicas. *Tesis de Pregrado*, 8-25. Lima, Perú. Obtenido de http://cybertesis.unmsm.edu.pe/bitstream/handle/20.500.12672/9114/Gonzalez_fd.pdf?sequence=1&isAllowed=y
- Graf, S., Ives, C., Rahman, N., & Ferri, A. (2011). AAT: a tool for accessing and analysing students' behaviour data in learning systems. *Proceedings of the 1st International Conference on Learning Analytics and Knowledge*, (págs. 174-179). doi:<https://doi.org/10.1145/2090116.2090145>
- Gutiérrez Braojos, C., Montejo Gamez, J., Marin Jimenez, A., & Campaña, J. (2018). Hybrid learning environment: Collaborative or competitive learning? *Virtual Reality*, 411-423. doi:10.1007/s10055-018-0358-z
- Hämäläinen, W., Laine, T., & Sutinen, E. (2006). Data Mining In Personalizing Distance Education Courses. *Computer Science*. doi:10.2495/1-84564-152-3/09



- Hand, D., & Adams, N. (2015). Data Mining. *Wiley StatsRef*, 1-7. doi:<https://doi.org/10.1002/9781118445112.stat06466.pub2>
- Highcharts Team. (2020). *highcharts.com*. Obtenido de <https://www.highcharts.com/>
- Hui, H., Ding, Y., Shi, Q., Li, F., Song, Y., & Yan, J. (Enero de 2020). 5G network-based Internet of Things for demand response in smart grid: A survey on application potential. *Applied Energy*, CCLVII. doi:<https://doi.org/10.1016/j.apenergy.2019.113972>
- IntelliBoard Team. (2020). *intelliboard.net*. Obtenido de <https://intelliboard.net/>
- J. Watson, H., & H. Wixom, B. (17 de Septiembre de 2007). The Current State of Business Intelligence. *Computer*, 40, 96-99. doi:10.1109/MC.2007.331
- Jovanović, J., Gašević, D., Dawson, S., Pardo, A., & Mirriahi, N. (2017). Learning analytics to unveil learning strategies in a flipped classroom. *Internet and Higher Education*, 33, 74-85. doi:10.1016/j.iheduc.2017.02.001
- Juhaňák, L., Zounek, J., & Rohlíková, L. (2019). Using process mining to analyze students' quiz-taking behavior patterns in a learning management system. *Computers in Human Behavior*, XCII, 496-506. doi:<https://doi.org/10.1016/j.chb.2017.12.015>
- Khan, K., Rehman, S., Aziz, K., Fong, S., & Sarasvady, S. (2014). DBSCAN: Past, present and future. *The Fifth International Conference on the Applications of Digital Information and Web Technologies (ICADIWT 2014)*, (págs. 232-238). Bangalore. doi:10.1109/ICADIWT.2014.6814687
- KlassData Team. (2016). *moodle.org*. Obtenido de Moodle Plugin: https://moodle.org/plugins/local_smart_klass
- Kreulen, J., Cody, W., Krishna, V., & Spangler, W. (2002). The integration of business intelligence and knowledge management. *IBM Systems Journal*, 41, 697 - 713. doi:10.1147/sj.414.0697
- Lage, M., Platt, G., & Treglia, M. (2000). Inverting the Classroom: A Gateway to Creating an Inclusive Learning Environment. *The Journal of Economic Education*, 31(1), 30-43. doi:<http://dx.doi.org/10.2307/1183338>
- Laugwitz, B., Held, T., & Schrepp, M. (2008). Construction and Evaluation of a User Experience Questionnaire. *HCI and Usability for Education and Work*, 5298, 63-76.
- LearnerScript Team. (2020). *learnerscript.com*. Obtenido de <https://learnerscript.com/>
- Llorente Cejudo, M. (2007). Hacia el e-learning desde el software libre: Moodle como entorno virtual de formación al alcance de todos. *Revista científica iberoamericana de comunicación y educación*, XV(28), 197-202. Obtenido de <http://hdl.handle.net/11162/86199>



- Lowell Bishop, J., & Verleger, M. (2013). The Flipped Classroom: A Survey of the Research. *ASEE Annual Conference and Exposition, Conference Proceedings*. Obtenido de https://www.researchgate.net/publication/285935974_The_flipped_classroom_A_survey_of_the_research
- Luna, J., Castro, C., & Romero, C. (2017). MDM tool: A data mining framework integrated into Moodle. *Computer Applications in Engineering Education*, 25, 90-102. doi:10.1002/cae.21782
- M. Gillies, R., & Boyle, M. (2011). Teachers' reflections of cooperative learning (CL): a two-year follow-up. *Teaching Education*, 63-78. doi:http://dx.doi.org/10.1080/10476210.2010.538045
- M. Müller, R., & Joachim Lenz, H. (2013). *Business Intelligence*. Heidelberg. doi:https://doi.org/10.1007/978-3-642-35560-8
- Mahesh Kumar, K., & Mohan Reddy, R. (Octubre de 2016). A fast DBSCAN clustering algorithm by accelerating neighbor searching using Groups method. *Pattern Recognition, LVIII*, 39-48. doi:https://doi.org/10.1016/j.patcog.2016.03.008
- Maldonado Mahauad, J., Perez Sanagustín, M., Kizilcec, R., Morales, N., & Munoz Gama, J. (2017). Mining theory-based patterns from Big data: Identifying self-regulated learning strategies in Massive Open Online Courses. *Computers in Human Behavior, LXXX*, 179-196. doi:https://doi.org/10.1016/j.chb.2017.11.011
- Maldonado, J., Pérez Sanagustín, M., Bermeo, J., Muñoz, L., Pacheco, G., & Espinoza, I. (2017). Flipping the Classroom with MOOCs. A Pilot Study Exploring Differences between Self-Regulated Learners. *2017 Twelfth Latin American Conference on Learning Technologies (LACLO)*. La Plata. doi:10.1109/LACLO.2017.8120934
- Marutho, D., Hendra Handaka, S., Wijaya, E., & Muljono. (2018). The Determination of Cluster Number at k-Mean Using Elbow Method and Purity Evaluation on Headline News. *2018 International Seminar on Application for Technology of Information and Communication*, (págs. 533-538). Semarang. doi:10.1109/ISEMANTIC.2018.8549751
- Mazza, R., Bettoni, M., Faré, M., & Mazzola, L. (2012). MOCLog – Monitoring Online Courses with log data. *1st Moodle Research Conference*, (págs. 132-139). Obtenido de <https://research.moodle.org/54/1/17%20-%20Mazza%20-%20MOCLog%20-%20Monitoring%20Online%20Courses%20with%20log%20data.pdf>



- Mazzola, L., Nidola, M., Milani, C., & Mazza, R. (2010). *L'aggiornamento del tool di monitoraggio delle attività degli studenti: GISMO 2.0*. Bari, Italia. Obtenido de <https://repository.supsi.ch/9789/>
- Mohri, M., Rostamizadeh, A., & Talwalkar, A. (2018). *Foundations of Machine Learning* (Segunda ed.). Londres, Inglaterra: Massachusetts Institute of Technology. Obtenido de <https://www.jstor.org/stable/j.ctt5hhcw1>
- Moodle, E. d. (2020). *Moodle.org*. Obtenido de https://docs.moodle.org/310/en/Main_page
- Moore, J., & Churchward, M. (2010). *Moodle 1.9 Extension Development Customize and extend Moodle by using its robust plugin systems*. Birmingham: Packt Publishing Ltd. Obtenido de https://subscription.packtpub.com/book/hardware_and_creative/9781847194244/1/ch011v11sec01/understanding-the-stack
- Moreno Marcos, P. M., Muñoz Merino, P., Alario Hoyos, C., Estévez Ayres, I., & Delgado Kloos, C. (2018). Analysing the predictive power for anticipating assignment grades in a massive open online course. *Behaviour and Information Technology*, 1021-1036. doi:<https://doi.org/10.1080/0144929X.2018.1458904>
- Moreno Marcos, P. M., Muñoz Merino, P., Maldonado Mahauad, J., Pérez Sanagustín, M., Alario Hoyos, C., & Delgado Kloos, C. (2020). Temporal analysis for dropout prediction using self-regulated learning strategies in self-paced MOOCs. *Computers&Education*. doi:<https://doi.org/10.1016/j.compedu.2019.103728>
- Mostow, J., & Beck, J. (2006). Some useful tactics to modify, map and mine data from intelligent tutors. *Natural Language Engineering*, 195-208. doi:10.1017/S1351324906004153
- Mwalumbwe, I., & Mtebe, J. (2017). Using learning analytics to predict students' performance in moodle learning management system: a case of mbeya university of science and technology. *The Electronic Journal of Information Systems in Developing Countries*, 79(1), 1-13. doi:<https://doi.org/10.1002/j.1681-4835.2017.tb00577.x>
- MySQL Team. (2020). *mysql.com*. Obtenido de <https://www.mysql.com/>
- Pardo, A., Jovanovic, J., Mirriahi, N., Dawson, S., Martinez Maldonado, R., & Gašević, D. (2016). Generating actionable predictive models of academic performance. *International Learning Analytics & Knowledge Conference 2016 - University of Edinburgh, Edinburgh, United Kingdom*, (págs. 474-478). New York. doi:10.1145/2883851.2883870
- Pedraza Pérez, R., Romero, C., & Ventura, S. (2011). A Java Desktop Tool for Mining Moodle Data. *Conference: Proceedings of the 4th International Conference on Educational Data Mining*.



- Netherlands. Obtenido de <https://www.semanticscholar.org/paper/A-Java-Desktop-Tool-for-Mining-Moodle-Data-Perez-Romero/cb859692767c6fbd1074ffc11dc8e5f3843bbdf4>
- Peña Rivera, J., & Suárez Daza, J. (2008). *Utilización de infoemación histórica para desiciones empresariales*. Obtenido de <https://repository.javeriana.edu.co/bitstream/handle/10554/7497/Tesis204.pdf?sequence=1&isAllowed=y>
- Rabbany, R., Takaffoli, M., & Zaïane, O. (2011). Analyzing participation of students in online courses using social network analysis techniques. *CiteSeerX*, 21-30. Obtenido de <https://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.419.6625>
- Ramírez Donoso, L. A., Pérez Sanagustín, M., Neyem, A., & Rojas Riethmuller, J. (2015). Promoviendo la Colaboración Efectiva en MOOCs a través de Aplicaciones Móviles. *Congreso IEEE CHILECON 2015*. Santiago de Chile. Obtenido de https://www.researchgate.net/publication/280778018_Promoviendo_la_Colaboracion_Efectiva_en_MOOCs_a_traves_de_Aplicaciones_Moviles
- Ramos, S. (2011). *Microsoft Business Intelligence: vea el cubo medio lleno* (Vol. I). Alicante, España: SolidQTM Press. Obtenido de <http://www.solidq.com/ebs/Microsoft-Business-Intelligence-vea-el-cubo-medio-lleno.pdf>
- Retalis, S., Papasalouros, A., Psaromiligkos, Y., Siscos, S., & Kargidis, T. (2006). Towards Networked Learning Analytics – A concept and a tool. *Fifth International conference on networked learning*, (págs. 1-8). Lancaster. Obtenido de <https://www.lancaster.ac.uk/fss/organisations/netlc/past/nlc2006/abstracts/pdfs/P41%20Retalis.pdf>
- Rice, W. (2006). *A complete guide to successful learning using Moodle* (4 ed.). New York. Obtenido de http://ead.iesgo.edu.br/file.php/1/apostilas/Moodle_Sample_e-book.pdf
- Rice, W. (2015). *Moodle E-Learning Course Development Third Edition A complete guide to create and develop engaging e-learning courses with Moodle* (Tercera ed.). Birmingham: Packt Publishing Ltd. Obtenido de https://subscription.packtpub.com/book/web_development/9781788472197/1/ch01lv11sec14/the-moodle-architecture
- Roman, V. (12 de Junio de 2019). *Aprendizaje No Supervisado en Machine Learning: Agrupación*. Obtenido de medium: <https://medium.com/datos-y-ciencia/aprendizaje-no-supervisado-en-machine-learning-agrupaci%C3%B3n-bb8f25813edc>



- Romero, C., & Ventura, S. (2016). Educational data science in massive open online courses. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, VII(1). doi:<https://doi.org/10.1002/widm.1187>
- Romero, C., Ventura, S., & Garcia, E. (Agosto de 2008). Data Mining in Course Management Systems: Moodle Case Study and Tutorial. *Computers & Education*, 51(1), 368-384. doi:<https://doi.org/10.1016/j.compedu.2007.05.016>
- Ros Martínez, I. (2008). Moodle, la plataforma para la enseñanza y organización escolar. *Ikastorratza, e-Revista de Didáctica*. Obtenido de http://www.ehu.eus/ikastorratza/2_alea/moodle.pdf
- Sánchez, P., García Saiz, D., & Zorrilla, M. (2012). Software product line engineering for e-learning applications: A case study. *2012 International Symposium on Computers in Education (SIE)*, (págs. 1-6). Obtenido de <https://ieeexplore.ieee.org/abstract/document/6403212>
- Schmitt, M. (2018). *moodle.org*. Obtenido de Moodle Plugins: https://moodle.org/plugins/block_analytics_graphs
- Schwendimann, B., Rodríguez Triana, M. J., Vozniuk, A., Prieto, L., & Shirvani Boroujeni, M. (2017). Perceiving Learning at a Glance: A Systematic Literature Review of Learning Dashboard Research. *IEEE Transactions on Learning Technologies*, 30-41. doi:10.1109/TLT.2016.2599522
- Shaun Bake, R., & Salvador Inventado, P. (2014). Educational Data Mining and Learning Analytics. En J. Ari Larusson, & B. White, *Learning Analytics* (págs. 61-75). New York. doi:https://doi.org/10.1007/978-1-4614-3305-7_4
- Shi, C., Wei, B., Wei, S., Wang, W., Liu, H., & Liu, J. (2021). A quantitative discriminant method of elbow point for the optimal number of clusters in clustering algorithm. *EURASIP Journal on Wireless Communications and Networking*. doi:<https://doi.org/10.1186/s13638-021-01910-w>
- Silva Peñafiel, G. E. (Marzo de 2018). Análisis de metodologías para la implementación de un data warehouse aplicado a la toma de decisiones del instituto nacional de Patrimonio Cultural Regional 3. *Tesis de Postgrado*. Ambato, Tungurahua, Ecuador. Obtenido de <https://repositorio.pucesa.edu.ec/bitstream/123456789/2367/1/76540.pdf>
- Smith, V. (2013). Data Dashboard as Evaluation and Research Communication Tool. *New Directions for Evaluation*(140), 21-45. doi:<https://doi.org/10.1002/ev.20072>
- Smola, A. (2008). *Introduction to Machine Learning* (Primera ed., Vol. I). Cambridge University. Obtenido de <https://alex.smola.org/drafts/thebook.pdf>



- Song, Y., Rowen, W., Medsker, C., & Ewen, E. (4 de Junio de 2001). An Analysis of Many-to-Many Relationships Between Fact and Dimension Tables in Dimensional Modeling. *Proceedings of the International Workshop on Design and Management of Data Warehouses*, 39, 1-13. Obtenido de <http://ceur-ws.org/Vol-39/paper6.pdf?ref=driverlayer.com/web>
- Sonnenberg, C., & Bannert, M. (2015). Discovering the Effects of Metacognitive Prompts on the Sequential Structure of SRL-Processes Using Process Mining Techniques. *Journal of Learning Analytics*, 72-100. doi:10.18608/jla.2015.21.5
- Spanjers, I., Könings, K., Leppink, J., Verstegen, D., Jong, N., Czabanowska, K., & Merriënboer, J. (2015). The promised land of blended learning: Quizzes as a moderator. *ELSEVIER*, 59-74. doi:<https://doi.org/10.1016/j.edurev.2015.05.001>
- Štrobl, J., Piorecký, M., & Krajča, V. (2017). Methods for automatic estimation of the number of clusters for k-means algorithm used on eeg signal: feasibility study. *Clinician and Technology*, XLVII(3), 81-87. Obtenido de <https://ojs.cvut.cz/ojs/index.php/CTJ/article/view/4474>
- Temple, J. (26 de Febrero de 2020). *MIT Technology Review*. Obtenido de [technologyreview: https://www.technologyreview.com/10-breakthrough-technologies/2020/](https://www.technologyreview.com/10-breakthrough-technologies/2020/)
- Teoh, E., & Kidd, D. (Diciembre de 2017). Rage against the machine? Google's self-driving cars versus human drivers. *Journal of Safety Research*, LXIII, 57-60. doi:<https://doi.org/10.1016/j.jsr.2017.08.008>
- Tran, T., Drab, K., & Daszykowski, M. (2013). Revised DBSCAN algorithm to cluster data with dense adjacent clusters. *Chemometrics and Intelligent Laboratory Systems*, CXX, 92-96. doi:<https://doi.org/10.1016/j.chemolab.2012.11.006>
- Uvidia Fassler, M. I. (Febrero de 2016). Descubrimiento de conocimiento en base de datos para la toma de decisiones en la unidad de nivelación y admisión de la ESPOCH. *Tesis de Postgrado*, 7-34. Ambato, Tungurahua, Ecuador. Obtenido de <https://repositorio.pucesa.edu.ec/bitstream/123456789/1601/1/76134.pdf>
- Valenzuela Zambrano, B., & Pérez Villalobos, M. V. (2013). Aprendizaje autorregulado a través de la plataforma virtual Moodle. *Educ*, XVI(1), 66-79. Obtenido de <https://www.redalyc.org/pdf/834/83428614009.pdf>
- Valle Torre, M., Tan, E., & Hauff, C. (2020). edX log data analysis made easy: introducing ELAT: An open-source, privacy-aware and browser-based edX log data analysis tool. *Proceedings of the Tenth International Conference on Learning Analytics & Knowledge*, (págs. 502-511). doi:<https://doi.org/10.1145/3375462.3375510>



- Vazquez Brust, A. (2019). *BSCAN: Machine Learning para detectar centros de actividad urbana*. Obtenido de bitsandbricks: <https://bitsandbricks.github.io/post/dbscan-machine-learning-para-detectar-centros-de-actividad-urbana/>
- Vue.js Team. (2020). *The Progressive JavaScript Framework*. Obtenido de vuejs.org: <https://vuejs.org/>
- Wander Rodrigues, M., Isotani, S., & Zárata, L. (2018). Educational Data Mining: A review of evaluation process in the e-learning. *Telematics and Informatics*, XXXV(6), 1701-1717. doi:<https://doi.org/10.1016/j.tele.2018.04.015>
- Wang, F., & Hannafin, M. (2005). Design-based research and technology-enhanced learning environments. *Educational Technology Research and Development*, (págs. 5-23). doi:<https://doi.org/10.1007/BF02504682>
- Watson, H., Ariyachandra, T., & Matyska, R. (2001). Data Warehousing Stages of Growth. *Information Systems Management*, 18(3), 42-50. doi:<https://doi.org/10.1201/1078/43196.18.3.20010601/31289.6>
- Westbrook, V. (2006). The virtual learning future. *The virtual learning future*, 11(4), 471-482. doi:<https://doi.org/10.1080/13562510600874276>
- Wild, I. (2017). *Moodle 3.x Developer's Guide*. Birmingham: Packt Publishing Ltd. Obtenido de https://books.google.com.ec/books?id=EHg5DwAAQBAJ&printsec=copyright&redir_esc=y#v=onepage&q&f=false
- Xing, W., & Du, D. (2018). Dropout Prediction in MOOCs: Using Deep Learning for Personalized Intervention. *Journal of Educational Computing*, LVII(3), 1-24. doi:<https://doi.org/10.1177/0735633118757015>
- Yassine, S., Kadry, S., & Sicilia, M. A. (2016). A framework for learning analytics in moodle for assessing course outcomes. *2016 IEEE Global Engineering Education Conference (EDUCON)*, (págs. 261-266). Abu Dhabi. doi:10.1109/EDUCON.2016.7474563
- Yen-Ting Liu, D., Froissard, J., & Deborah, R. (2015). An enhanced learning analytics plugin for Moodle: student engagement and personalised intervention. *ascilite2015*, 180-189. Obtenido de <https://research.moodle.org/76/1/Liu%20%282015%29%20An%20enhanced%20learning%20analytics%20plugin%20for%20Moodle.pdf>
- Yen-Ting Liu, D., Froissard, J., & Richards, D. (2015). An enhanced learning analytics plugin for Moodle: student engagement and personalised intervention. *ASCILITE 2015*, (págs. 180-189). Perth. Obtenido de



<https://research.moodle.org/76/1/Liu%20%282015%29%20An%20enhanced%20learning%20analytics%20plugin%20for%20Moodle.pdf>

- Zafra, A., Romero, C., & Ventura, S. (2013). DRAL: a tool for discovering relevant e-activities for learners. *Knowledge and Information Systems*, 211-250. doi:<https://doi.org/10.1007/s10115-012-0531-8>
- Zhang, X.-D. (2020). Machine Learning. En *Matrix Algebra Approach to Artificial Intelligence* (págs. 223-440). Singapore. doi:https://doi.org/10.1007/978-981-15-2770-8_6
- Zins, C. (2007). Conceptual Approaches for Defining Data, Information,. *Journal of the American Society for Information Science and Technology*, 58, 479-493. doi:<https://doi.org/10.1002/asi.20508>
- Zorrilla, M., & García Saiz, D. (Abril de 2013). A service oriented architecture to provide data mining services for non-expert data miners. *Decision Support Systems*, LV(1), 399-411. doi:<https://doi.org/10.1016/j.dss.2012.05.045>
- Zwart, H. (2015). Human Genome Project: History and Assessment. En H. Zwart, *International Encyclopedia of Social & Behavioral Sciences* (Segunda ed., Vol. XI, págs. 311-317). doi:10.1016/B978-0-08-097086-8.82036-X



Anexos

Anexo 1 Certificado de presentación del artículo de posición

Figura 1

Certificado obtenido en LACLO 2020



Nota. Obtenido de LACLO 2020



Proposal of Design and Evaluation of a Dashboard for the Analysis of Learner Behavior and Prediction of Dropout in Moodle

Sigua L. Edisson

Departamento de Ciencias de la Computación
Universidad de Cuenca
Cuenca, Ecuador
edisson.sigua1407@ucuenca.edu.ec

Pesantez C. Paola

Departamento de Ciencias de la Computación
Universidad de Cuenca
Cuenca, Ecuador
paola.pesantez@ucuenca.edu.ec

Aguilar Y. Bryan

Departamento de Ciencias de la Computación
Universidad de Cuenca
Cuenca, Ecuador
bryan.aguilar@ucuenca.edu.ec

Maldonado M. Jorge

Departamento de Ciencias de la Computación
Universidad de Cuenca
Cuenca, Ecuador
jorge.maldonado@ucuenca.edu.ec

Abstract— The rapid development of technology has meant that over the past two decades Information and Communications Technologies (ICT) become increasingly involved in the teaching process and seek to change traditional learning models. With the support of modern technology, virtual platforms that encourage the adoption of a new learning paradigm in which geographical/temporal limitations no longer pose a difficulty have been developed and refined. These virtual learning platforms, also known as Learning Management Systems (LMS), store student and teacher interactions with course resources, and these interactions are stored in database engines. However, all the information generated by LMS has not been processed in a way that is helpful for the use of teachers and students, mainly because in most cases, students' interactions with these systems focus on downloading class material, delivering assignments, and reading announcements, leaving aside indicators that can be presented in the form of visualizations that allow actions to be taken during the development of the learning process. Thus, this study proposes the implementation and evaluation of a dashboard for the analysis of learner behavior and prediction of dropout on the Moodle platform. The proposed tool will help students to manage their learning process, easily and effectively monitor their progress in an online course, and teachers to know what students do before, during and after a virtual class. The latter for the purpose of being able to detect early students at risk of dropping out.

Keywords— *Learning Analytics, Dashboard, Moodle, Dropout, Prediction.*

I. INTRODUCCIÓN

El desarrollo tecnológico ha sido uno de los pilares fundamentales en la evolución humana hacia lo que somos hoy en día. La razón principal de esta relación de desarrollo mutuo se basa en el hecho de que todo proceso evolutivo es una combinación de componentes y/o procesos para cubrir algún propósito en un área humana específica [1]. Algunas de las áreas que más se han beneficiado del desarrollo tecnológico son la computación, medicina, robótica, electricidad y telecomunicaciones [2]. Actualmente, podemos establecer una conversación en tiempo real con otra persona en cualquier parte del mundo y por medio de las redes sociales podemos acceder a información de cualquier acontecimiento en instantes. Con los avances logrados en Inteligencia

Artificial (AI – *Artificial Intelligence*) se han empezado a construir vehículos con sistemas de manejo autónomos capaces de llegar a un destino sin necesidad de interacción humana [3]. También gracias a la tecnología, la medicina ha tenido grandes avances como el mapeo del Genoma Humano, el mismo que se pudo estudiar sistemáticamente con la ayuda de programas informáticos para la obtención de patrones y asociaciones [4]. Además, a corto plazo, el número de dispositivos conectados a Internet aumentará de manera exponencial gracias a la velocidad de transferencia rápida, alta confiabilidad, seguridad robusta, bajo consumo de energía y gran cantidad de conexiones de las redes 5G [5]. De esta forma es posible ver como la tecnología tiene influencia en todas las áreas humanas.

En el caso específico del área de la educación, los avances en las Tecnologías de Información y Comunicación (ICT – *Information and Communications Technology*) han tratado de contribuir al enfoque tradicional de aprendizaje presencial, convirtiéndose en parte integral de los procesos de enseñanza y aprendizaje [6] y generando áreas de interés investigativo tales como las Analíticas del Aprendizaje (LA – *Learning Analytics*) y Minería de Datos Educativos (EDM – *Educational Data Mining*) [7]. Gracias al uso de las tecnologías, por medio de la web y otros canales, el sector educativo ha podido superar las limitaciones geográficas/temporales que los esquemas tradicionales de enseñanza-aprendizaje conllevan, dando lugar así al nacimiento de nuevos modelos de enseñanza y aprendizaje como e-formación, e-learning o b-learning [8]. De los tres modelos antes mencionados, el que incorpora significativamente las ICT a su proceso de enseñanza es el Aprendizaje Combinado (b-learning – *Blended Learning*), el cual se define como un enfoque que combina los beneficios ofrecidos por los componentes de aprendizaje presencial y en línea [9]. Como resultado, en los últimos años ha aumentado el interés en el diseño de entornos de aprendizaje híbridos, ya que combinar las actividades de enseñanza presencial y en línea ofrece nuevas oportunidades para potenciar y optimizar el aprendizaje [10]. De hecho, el aprendizaje híbrido, respaldado por plataformas virtuales, se está convirtiendo en una de las perspectivas de aprendizaje y enseñanza más impactantes en la educación superior [11].



Anexo 3. Relaciones de tablas de la base de datos de Moodle, identificando tablas relevantes y atributos de interés para el estudio

Figura 1

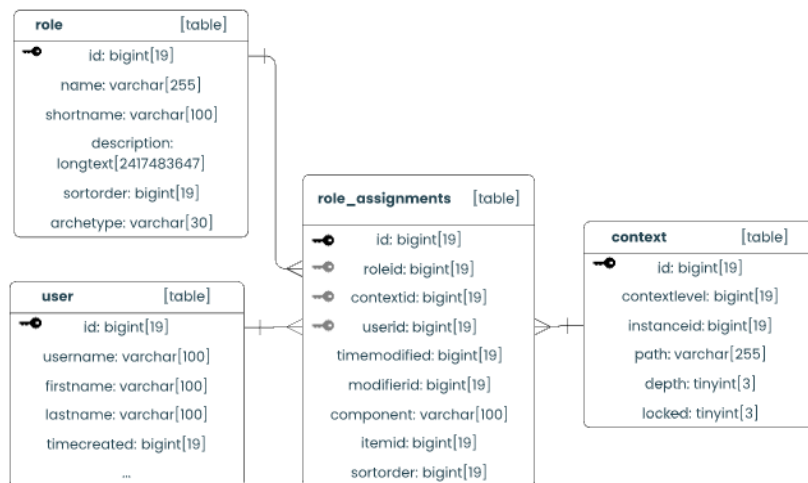
Tabla del curso con sus atributos más relevantes

course	[table]
id: bigint[19]	
fullname: varchar[254]	
shortname: varchar[255]	
startdate: bigint[19]	
enddate: bigint[19]	
timecreated: bigint[19]	
...	

Nota. Se ha extraído una representación simplificada de esta tabla indicando los atributos más importantes para el estudio. Elaboración Propia.

Figura 2

Relaciones entre tablas que permiten determinar qué rol tiene un usuario en el sistema

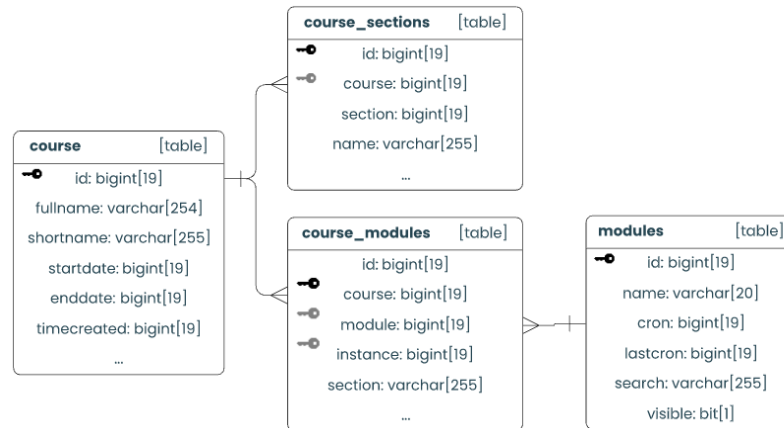


Nota. Se ha extraído una representación simplificada de este conjunto de tablas indicando los atributos más importantes para el estudio. Elaboración Propia.



Figura 3

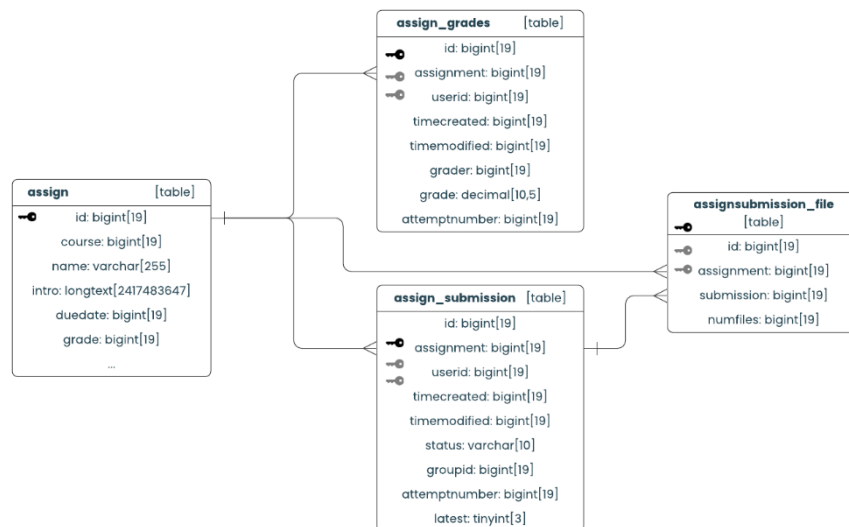
Relaciones entre tablas que permiten conocer las secciones y módulos en un curso



Nota. Se ha extraído una representación simplificada de este conjunto de tablas indicando los atributos más importantes para el estudio. Elaboración Propia.

Figura 4

Relaciones entre tablas que permiten conocer aspectos referentes a las tareas (assign)

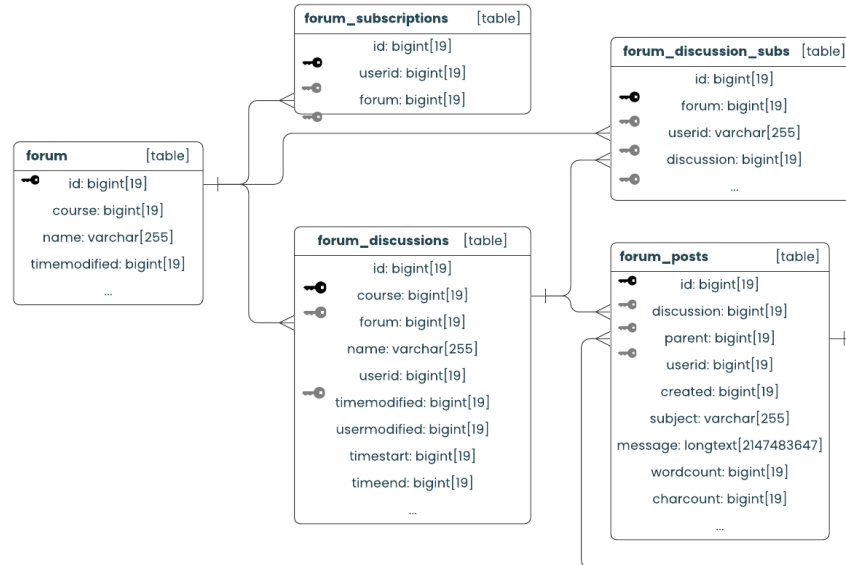


Nota. Se ha extraído una representación simplificada de este conjunto de tablas indicando los atributos más importantes para el estudio. Elaboración Propia.



Figura 5

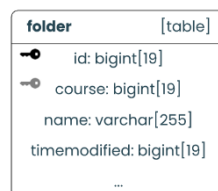
Relaciones entre tablas que permiten conocer aspectos referentes a los foros (forum)



Nota. Se ha extraído una representación simplificada de este conjunto de tablas indicando los atributos más importantes para el estudio. Elaboración Propia.

Figura 6

Tabla representa la actividad de tipo carpeta con sus atributos más relevantes

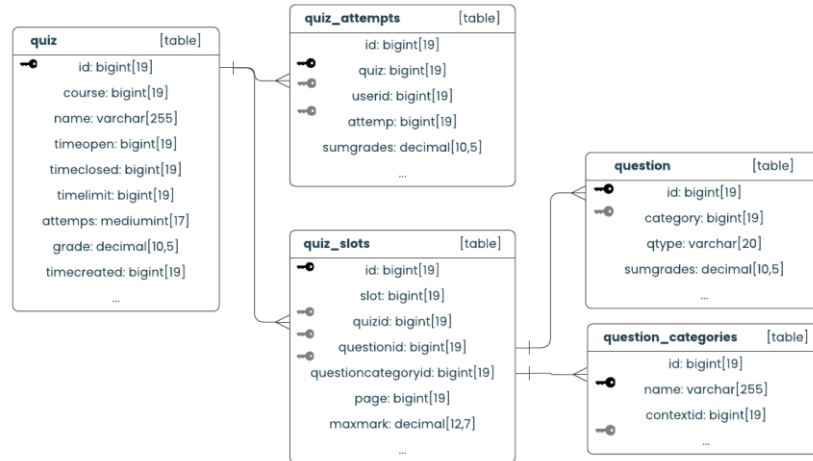


Nota. Se ha extraído una representación simplificada de esta tabla indicando los atributos más importantes para el estudio. Elaboración Propia.



Figura 7

Relaciones entre tablas que permiten conocer aspectos referentes a las evaluaciones (quiz)



Nota. Se ha extraído una representación simplificada de este conjunto de tablas indicando los atributos más importantes para el estudio. Elaboración Propia.

Figura 8

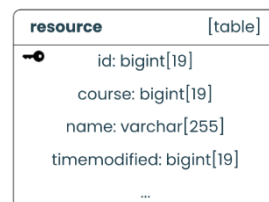
Tabla que Representa la Actividad de Tipo Página con sus atributos más relevantes



Nota. Se ha extraído una representación simplificada de esta tabla indicando los atributos más importantes para el estudio. Elaboración Propia.

Figura 9

Tabla Representa la Actividad de tipo Recurso con sus atributos más relevantes



Nota. Se ha extraído una representación simplificada de esta tabla indicando los atributos más importantes para el estudio. Elaboración Propia.



Figura 10

Tabla Representa el Recurso de Tipo url con sus atributos más relevantes

url [table]	
id: bigint[19]	
course: bigint[19]	
name: varchar[255]	
timemodified: bigint[19]	
...	

Nota. Se ha extraído una representación simplificada de esta tabla indicando los atributos más importantes para el estudio. Elaboración Propia.

Figura 11

Tabla que almacena las Calificaciones de los Estudiantes en las diferentes Actividades

grade_grades [table]	
id: bigint[19]	
itemid: bigint[19]	
userid: bigint[19]	
rawgrademax: decimal[10,5]	
rawgrademin: decimal[10,5]	
finalgrade: decimal[10,5]	
timecreated: bigint[19]	
timemodified: bigint[19]	
...	

Nota. Se ha extraído una representación simplificada de esta tabla indicando los atributos más importantes para el estudio. Elaboración Propia.



Figura 12

Tabla que Almacena los ítems Calificables de los Estudiantes

grade_items [table]	
id: bigint[19]	
courseid: bigint[19]	
itemname: varchar[255]	
itemtype: varchar[30]	
itemmodule: varchar[30]	
iteminstance: bigint[19]	
grademax: decimal[10,5]	
grademin: decimal[10,5]	
gradepass: decimal[10,5]	
timecreated: bigint[19]	
timemodified: bigint[19]	
...	

Nota. Se ha extraído una representación simplificada de esta tabla indicando los atributos más importantes para el estudio. Elaboración Propia.

Figura 13

Tabla de Interacciones de los Estudiantes o Docentes con las Actividades del Curso

logstore_standard_log [table]	
id: bigint[19]	
eventname: varchar[255]	
component: varchar[100]	
action: varchar[100]	
target: varchar[100]	
objecttable: varchar[50]	
objectid: bigint[19]	
contextid: bigint[19]	
userid: bigint[19]	
courseid: bigint[19]	
relateduserid: bigint[19]	
timecreated: bigint[19]	
...	

Nota. Se ha extraído una representación simplificada de esta tabla indicando los atributos más importantes para el estudio. Elaboración Propia.